TOMO II

Juan Mejía Trejo

EVALUACIÓN DE IMPACTO SOCIAL EN PROYECTOS DE INNOVACIÓN VÍA STATA

Métodos: Estimación por Variable Instrumental, Regresión Discontinua, Elección de método a usar



TOMO II

Juan Mejía Trejo

EVALUACIÓN DE IMPACTO SOCIAL EN PROYECTOS DE INNOVACIÓN VÍA STATA

Métodos: Estimación por Variable Instrumental, Regresión Discontinua, Elección de método a usar



Este libro fue financiado con el fondo federal PROINPEP, Programa de Incorporación y Permanencia de Posgrado en el Programa Nacional de Posgrado de Calidad del CONACYT.

Este libro fue sometido a un proceso de dictamen por pares de acuerdo con las normas establecidas por el Comité Editorial del Centro Universitario de Ciencias Económico Administrativas (CUCEA) de la Universidad de Guadalajara (UdeG).

Primera edición, 2021

Centro Universitario de Ciencias Económico Administrativas
Periférico Norte No. 799
Núcleo Universitario, Los Belenes,
45100 Zapopan, Jalisco

ISBN Tomo: 978-607-571-167-6
ISBN Colección: 978-607-571-165-2
https://doi.org/10.55965/abib.9786075711676.2021b

Hecho en México
Made in Mexico

Índice

Introducción

En el mundo empresarial e industrial, el diseño e implementación de proyectos normalmente toma en cuenta el impacto económico, financiero y hasta el político o ambiental en la introducción de innovaciones. Sin embargo, en los primeros veinte años del siglo XXI, han sucedido diversos acontecimientos que han demostrado que la evaluación de impacto social en la generación de bienestar, es de vital importancia. De hecho, en los tiempos de la nueva normalidad que se avizora como la era PosCOVID-19, esto toma particular relevancia dado que todas las políticas y acciones que emitan empresas y gobiernos, deberán contar con el aval necesario de una evaluación de impacto social a la introducción de innovaciones.

Es por esta razón, que la obra: *EVALUACIÓN DE IMPACTO SOCIAL EN PROYECTOS DE INNOVACIÓN VÍA STATA. Métodos: Estimación por Variable Instrumental, Regresión Discontinua, Elección de método a usar TOMO II*, está orientada a describir tanto a propios como ajenos al tema, lo qué es el impacto social, sus características, condiciones e implicaciones, los principales métodos utilizados para calcularla así como las oportunidades que se vislumbran en la era PosCOVID-19, que demanda que los recursos y acciones de innovación a diseñar e implementar, relfejen altos estándares de impacto social que fomenten el bienestar, particularmente en los países emergentes.

Para logralo, esta obra está dividida en una colección de dos tomos, correspondiendo al Tomo II:

Capítulo 6. Estimación por Variable Instrumental (*VI. Instrumental Variable*). En este capítulo reporta al lector los tipos de estimaciones, el enfoque de mínimos cuadrados de dos etapas, lo que implica el cumplimiento imperfecto y los instrumentos débiles en los estimados, lo que significa el efecto de tratamiento marginal y lo que significa la promo-

ción aleatoria en la técnica, lista de verificación. Cierra el capítulo con los alcances y limitaciones, así como con un ejemplo de aplicación en STATA.

Capítulo 7. Regresión Discontínua (*RD. Regression Discontinuity*). El capítulo está diseñado para presentar la teoría de la regresión de discontinuidad, tanto nítida como difusa, los pasos que conllevan a su aplicación, posibles variaciones, la verificación de la validez del diseño, ventajas y desventajas de la técnica, comparaciones con otras (*pipeline*), limitaciones y alcances de la técnica, lista de verificación. Se incluye un ejemplo utilizando STATA.

Capítulo 8. Elección del método a utilizar. Dada la variedad de técnicas a utilizar y con el fin de ayudar al lector a identificar la más adecuada, este capítulo está diseñado para explicar la importancia de determinar los grupos de comparación, la implicación de lo que es la prospectiva y cómo crear un grupo de comparación, la identificación de los beneficiarios y su priorización. Lo cual, sirve para dar base a cómo comparar y respaldar a los métodos de evaluación de impacto, lo que significa determinar la unidad de intervención mínima, cómo lograrlo y evitar sus dificultades de implementación. Dada la intervención humana, se percibe cómo interviene la conducta y sus efectos no intencionados, la imperfección del cumplimiento, los efectos indirectos o de derrame (*spillover*) y el desgaste en la muestra. Se aportan sugerencias de diseño así como el tomar en cuenta los efectos y su persistencia en el tiempo. Asimismo, se explica la posibilidad de combinar diversas opciones de tratamiento en la evaluación de impacto en diferentes niveles de tratamiento así como la adaptación de evaluación de intervenciones múltiples.

Capítulo 9. Elección de la muestra y diseño de recolección de datos. Con el fin de apoyar a un prospecto de proyecto, se dan aquí los elementos para determinar la muestra a partir de los métodos de muestreo más conocidos, lo que implica el cálculo de potencia de muestra así como los grupos de tratamiento vs comparación. Se citan los errores potenciales que se deben evitar, se define qué es y cómo calcular la potencia estadística, el significado de los clúster relacionado a la asignación aleatoria, recomendaciones para recopilar datos y cómo hacerlos indicadores que midan.

Capítulo 10. Lineamientos de gestión de proyectos de evaluación de impacto. En este capítulo, se resalta la importancia de los proyectos de innovación en el impacto social, cómo la sociedad se involucra y hace más participativa en las etapas de proyectos de innovación, se especifica a la empresa y su compromiso social, se plantean en general, objetivos de desarrollo sostenible de la ONU y el capital comunitario. Con esto de referencia, se especifican los componentes del equipo de gestión de la evaluación de impacto, sugiriendo cómo establecer colaboraciones y determinando tanto un cronograma como un presupuesto a un proyecto de gestión de proyectos de evaluación de impacto.

Capítulo 11. Evaluación de impacto social en la era PosCOVID-19. Finalmente, se cierra la obra con la descripción general del fenómeno COVID-19 a nivel mundial así como su impacto económico, comercio internacional e internet. Para América Latina y el Caribe, se describen las oportunidades de la evaluación de impacto social en la introducción de innovaciones en los sistemas de salud, educación, empleo y pobreza, en la mipyme. Se concluye con un reporte sobre futuro mundial posCOVID-19, el impacto social del COVID-19 en México así como la estructura operativa estrategias ante COVID-19, que son las bases para la introducción de evaluación de impacto social en la introducción de innovaciones en la era PosCOVID-19.

CAPÍTULO 6.

Estimación por Variable Instrumental (*IV. Instrumental Variables*)

En el análisis de la asignación aleatoria del capítulo 3, se asume que el administrador del programa tiene la facultad para asignar la intervención a los grupos de tratamiento y de comparación, y que los asignados al tratamiento participan en el programa y los asignados al grupo de comparación no lo hacen. El pleno cumplimiento se logra con mayor frecuencia en pruebas de laboratorio o en ensayos médicos, donde el investigador puede asegurar, primero, que todos los sujetos del grupo de tratamiento reciban un determinado tratamiento y, segundo, que no lo reciba ninguno de los sujetos del grupo de comparación. De manera más general, en el capítulo 3 se asume que los programas pueden determinar quiénes son los participantes potenciales, excluyendo a algunos y asegurando que otros participen a toda la población elegible, de modo que no sería ético asignar aleatoriamente a las personas a los grupos de tratamiento y de comparación, y excluir a participantes potenciales en aras de la evaluación. Por ende, se necesita una manera alternativa de evaluar el impacto de este tipo de programas.

El método denominado variables instrumentales (**IV.** *Instrumental Variables)* puede resultar útil para evaluar los programas con cumplimiento imperfecto, inscripción voluntaria o cobertura universal. En general, para estimar los impactos, el método **IV** depende de una fuente externa de variación para determinar el estatus del tratamiento. El método puede aplicarse a un amplio espectro de situaciones, más allá de la evaluación de impacto. Se puede pensar en una **IV** como algo que escapa al control del individuo y que influye en su probabilidad de participar en un programa pero que, de otra manera, no está asociado con las características de dicho individuo.

Esta variación externa, o **IV**, puede ser generada por las reglas de funcionamiento del programa que está bajo el control de los encargados del mismo o de los equipos de evaluación. Para producir evaluaciones de impacto válidas, esta fuente externa de variación, o **IV**, debe satisfacer un cierto número de condiciones. Se ha observado que la asignación aleatoria del tratamiento, como se señaló en el capítulo 3, es un muy buen instrumento y que satisface las condiciones necesarias. **(Gertler et al., 2017):**

- Primero, se usará como una extensión del método de asignación aleatoria cuando no todas las unidades cumplen con su asignación de grupo.
- En segundo lugar, se recurrirá a él para diseñar una promoción aleatoria del tratamiento, un método de evaluación que puede funcionar en algunos programas que ofrecen inscripción voluntaria o cobertura universal.

Ahora se tratará de los métodos, que *relajan* la suposición de exogeneidad de los **OLS** o **PSM** y que también, son robustos a *sesgos de selección que varían en el tiempo*, a diferencia del método **DD.** Recuerde que para los métodos **DD,** no es posibel controlar el sesgo de selección que cambia con el tiempo (Capítulo 5). Al *relajar* el *supuesto de exogeneidad*, el método **IV** hace diferentes supuestos de identificación, de los métodos anteriores, aunque los supuestos subyacentes del **IV** pueden no aplicarse en todos los contextos. Recuerde la configuración discutida en el capítulo 2 de una ecuación de estimación que compara los resultados de los *grupos tratados y no tratados* **(Khandker et al., 2017)**:

$$Y_i = \alpha X_i + \beta T_i \quad \text{.................................} \quad (6.1)$$

Si la asignación de tratamiento T, es aleatoria en la **ecuación 6.1**, el *sesgo de selección* no es un problema a nivel de aleatorización (ver capítulo 3). Sin embargo, la asignación de tratamiento puede no ser aleatorio debido a dos factores generales:

- Primero, puede existir *endogeneidad* en la orientación o colocación del programa, es decir, los programas se colocan deliberadamente en áreas que tienen características específicas (como oportunidades de ingresos o normas sociales) que pueden o no ser observado y eso también está correlacionado con los resultados $\boldsymbol{Y}$.
- Segundo, la *heterogeneidad no observada individual* derivada de la autoselección de beneficiarios individuales en el programa también se implican en una configuración experimental.

Como se discutió en el capítulo 2, el *sesgo de selección* puede ser el resultado de ambos factores, porque las características no observadas en el término de error, contendrán variables que también se correlacionan con el tratamiento ficticio (*dummy*) de $\boldsymbol{T}$. Es decir, $\mathbf{cov}\ (\boldsymbol{T}, \boldsymbol{\varepsilon}) <> \mathbf{0}$ implica violación de uno de los supuestos clave de la **OLS** en la obtención estimaciones no sesgadas: *independencia de los regresores del término de perturbación* $\boldsymbol{\varepsilon}$. La correlación entre $\boldsymbol{T}$ y $\boldsymbol{\varepsilon}$, sesga naturalmente las otras estimaciones en la ecuación, incluyendo la estimación del efecto del programa $\boldsymbol{\beta}$.

La **ecuación 6.1**, así como las preocupaciones correspondientes sobre la *endogeneidad*, pueden generalizarse a una configuración de *panel*. En este caso, las características no observadas a lo largo del tiempo, pueden estar correlacionadas con el programa y otros con covariables observadas. Hasta cierto punto, este problema se discutió en el capítulo 5. Los métodos **DD** resolvieron el problema suponiendo que *no se observan* las características de las unidades dirigidas y no dirigidas, eran invariantes en el tiempo y luego por diferenciación, de la heterogeneidad. Cuando los datos de panel están disponibles, los métodos **IV** permiten un mayor vista matizada de la heterogeneidad no observada, lo que permite que estos factores cambien con el tiempo (como el talento empresarial no observado de los sujetos seleccionados, la capacidad de mantener relaciones sociales vínculos y redes de innovación, etc., todo lo cual puede variar con la duración del programa).

El **IV** tiene como objetivo limpiar la correlación entre $\boldsymbol{T}$ y $\boldsymbol{\varepsilon}$. Es decir, la variación en $\boldsymbol{T}$ que no está correlacionado con $\boldsymbol{\varepsilon}$ necesita ser aislado. Para

hacerlo, uno necesita encontrar un instrumento variable, denotada ***Z***, que satisface las siguientes condiciones:

- Correlacionado con ***T: cov (Z,T) <>0***
- Uncorrelacionado con ***ε: cov (Z, ε) = 0***

Por lo tanto, el instrumento **Z** afecta la selección en el programa, pero no está correlacionado con factores que afectan los resultados (también conocido como *restricción de exclusión*). Un problema relacionado, es que el *error de medición* en la participación observada, puede subestimar o sobreestimr el impacto del programa. Como se discutió en el capítulo 3, un **IV** puede ser introducido para resolver esta atenuación de sesgo mediante el cálculo de una intención de trato (**ITT**. *Intention to Treat*) o estimación del programa. Esta estimación, explicaría que la participación real, sea diferente de la participación prevista debido a las reglas de selección y elegibilidad. **Khandker (2006)** proporciona un ejemplo de cómo las preocupaciones con respecto a la *exogeneidad* y la *atenuación de sesgo* pueden ser abordadas.

Tipos de estimaciones

Una evaluación de impacto siempre estima el impacto de un programa comparando los resultados de un grupo de tratamiento con la estimación de un contrafactual obtenido de un grupo de comparación válido. En el capítulo 3 se asumía que había pleno cumplimiento en el tratamiento, es decir, que todas las unidades a las que se ofrecía un programa se inscribían en él y que ninguna de las unidades del grupo de comparación recibía el programa. En este escenario, se estimaba el efecto promedio del tratamiento para la población. En la evaluación de los programas en el mundo real, donde los participantes potenciales pueden decidir si se inscriben o no, el pleno cumplimiento es menos común que en contextos como los experimentos de laboratorio. En la práctica, los programas suelen ofrecer tratamiento a un grupo específico, y algunas unidades participan y otras no. En este caso, sin

pleno cumplimiento, las evaluaciones de impacto pueden estimar el efecto de ofrecer un programa o el efecto de participar en el programa.

La *intención de tratar* (**ITT.** *Intention To Treat*) es un promedio ponderado de los resultados de los participantes y no participantes en el grupo de tratamiento versus el resultado promedio del grupo de comparación. Es importante en aquellos casos en los que se intenta determinar el impacto promedio de ofrecer un programa y la inscripción en el grupo de tratamiento es voluntaria. En cambio, es posible que también se desee conocer el impacto de un programa en el grupo de individuos a los que se ofrece el programa y que realmente participan. La **ITT** estima la diferencia en los resultados entre las unidades asignadas al grupo de tratamiento y las unidades asignadas al grupo de comparación, independientemente de si las unidades asignadas al grupo de tratamiento reciben en efecto el tratamiento. Este impacto estimado se denomina tratamiento en los tratados (**TOT**. *Treatment On The Treated*) estima la diferencia en los resultados entre las unidades que realmente reciben el tratamiento y el grupo de comparación. La **ITT** y el **TOT** serán iguales cuando haya pleno cumplimiento **(Gertler et al., 2017).**

Enfoque de mínimos cuadrados de dos etapas en IV

Para aislar la parte de la variable de tratamiento, que es independiente de otra de características no observadaa que afectan el resultado, primero se hace la regresión al grupo de tratamiento, con el instrumento ***Z*,** la que covaría en la **ecuación 6.1** un disturbio, u_i. Este proceso se conoce como primera etapa de regresión (*first-stage regression*) **(Khandker et al., 2017)**:

$$T_i = \gamma Z_i + \Phi X_i + u_i \ldots\ldots\ldots\ldots (6.2)$$

El tratamiento predicho en esta regresión, ***T***, por lo tanto, refleja la parte de tratamiento afectado solamente por ***Z*** y por lo tanto, implica solamente

la variación exógena en el tratamiento. ***T*** es entonces sustituída para la **ecuación 6.1** de tratamiento para crear la siguiente forma reducida de la regresión:

$$Y_i = \alpha X_i + \beta(\gamma' Z_i + \Phi' X_i + u_i) + \varepsilon_i \ldots\ldots\ldots\ldots\ldots(6.3)$$

El método **IV** (también conocido como mínimos cuadrados de dos etapas (**2SLS**. *Two-Stage Least Squares*) estimado del impacto del programa, es

entonces β_{iv} **.** Específicamente, observando en $\boldsymbol{Y_i = \beta T_i + \varepsilon_i}$, una versión simplificada de la ecuación **6.1**, y teniendo en cuenta el supuesto de **cov (*Z*, *ε*) = 0,** se puede escribir el efecto del tratamiento bajo **IV (**β**)**, como **cov (*Y*,*Z*) / cov (*T*,*Z*):**

$$\text{cov}(Y_i, Z_i) = \text{cov}[(\beta T_i + \varepsilon_i), Z_i] = \beta\, cov(T_i, Z_i) \ldots\ldots\ldots\ldots\ldots(6.4)$$

$$\Rightarrow \frac{\text{cov}(Y_i, Z_i)}{\text{cov}(T_i, Z_i)} = \beta. \ldots\ldots\ldots\ldots\ldots(6.5)$$

Esta derivación, se vuelve importante al examinar los efectos de la calidad del instrumento sobre el impacto estimado del programa bajo el método **IV** (ver la siguiente sección). A través de la instrumentación, por lo tanto, ***T** se limpia* de su correlación, con el *término de error*. Si los supuestos **cov (*T*, *Z*) <> 0 y cov (*Z*, *ε*) = 0** se mantienen, entonces el método **IV** identifica consistentemente el impacto medio del programa, atribuible al instrumento. Específicamente, se puede demostrar que $\boldsymbol{\beta_{IV} = \beta}$ **+ cov (*Z*, *ε*) / cov (*Z*, *T*).** La idea se discute más adelante.

Aunque la información detallada sobre la implementación y participación del programa, puede revelar directamente la presencia de *selección de sesgo*, la *endogeneidad* del tratamiento también puede ser evaluado con la *prueba de Wu-Hausman*, el cual consiste en:

- Primero, haga regresión ***T*** en ***Z*** y de las otras covariables exógenas ***X***, y obtenga los residuos u_i***'***. Estos residuos reflejan *toda la heterogeneidad* no observada, que afecta al tratamiento no capturado por los instrumentos y las variables exógenas en el modelo.
- Haga regresión de ***Y*** en ***X***, ***Z*** y u_i***'***. Si el coeficiente de u_i***'***, es estadísticamente diferente de cero, las características no observadas que afectan conjuntamente el tratamiento ***T*** y los resultados ***Y*** son significantes y se rechaza la nulidad de que ***T*** sea exógeno.

El modelo **IV** tiene algunas variaciones. Por ejemplo, uno podría reescribir la ecuación instrumento, como un *modelo de respuesta binaria no lineal* (como un *probit* o *logit)* y usar la propensión de puntaje previsto, como el método **IV** para la colocación del programa. Además, si existen datos de panel, el método **IV** puede combinarse con el enfoque de *panel efectos fijos,* como sigue **(Semykina y Wooldridge, 2005):**

$$Y_{it} = \delta Q_{it} + \eta_i + v_{it}\ , t = 1,\ldots,T,\ldots\ldots\ldots\ldots\ldots(6.6)$$

En la **ecuación 6.6**, η_i es el *efecto fijo no observado* (discutido en el capítulo 5) que puede ser correlacionado con la participación en el programa, Q_{it} representa un error idiosincrásico variante en el tiempo, y Q_{it} es un vector de covariables que incluye variables exógenas ***X***, así como el programa ***T***. En esta especificación, por lo tanto, la correlación entre η_i y la variable de tratamiento en Q_{it} se contabiliza a través del *enfoque de efectos* fijos o de *diferenciación*. Se introducen instrumentos Z_{it} para permitir la correlación entre algunos de los regresores en Q_{it} (como ***T***) y v_{it}. La idea aquí, sería encontrar instrumentos correlacionados con captación del programa (pero no de resultados) con el tiempo. Los supuestos restantes y la interpretación de la estimación son similares.

El cumplimiento imperfecto

Como ya se ha señalado, en los programas sociales del mundo real, el pleno cumplimiento con los criterios de selección de un programa (y, por ende, la adhesión a la condición de tratamiento o comparación) es deseable, y los responsables de las políticas y los equipos de evaluación por igual suelen intentar acercarse lo más posible a ese ideal. Sin embargo, en la práctica, no siempre se consigue un cumplimiento del **100%** de las asignaciones a los grupos de tratamiento y comparación, a pesar de los esfuerzos del encargado del programa y del equipo de evaluación. Para empezar, hay que subrayar que la mejor solución para el cumplimiento imperfecto consiste sencillamente en evitarlo. En este sentido, los administradores del programa y los responsables de las políticas deberían intentar que el cumplimiento sea lo más alto posible en el grupo de tratamiento y lo más bajo posible en el grupo de comparación.

En circunstancias de no cumplimiento, una segunda opción consiste en estimar lo que se conoce como el efecto local promedio del tratamiento (**LATE.** *Local Average Treatment Effec*t). El **LATE** debe ser interpretado con cuidado, ya que representa los efectos del programa solo para un subgrupo específico de la población. En particular, cuando hay incumplimiento en el grupo de tratamiento y en el de comparación, el **LATE** es el impacto en el subgrupo de cumplidores. En el ejemplo de la formación docente, si hay incumplimiento en ambos grupos, la estimación **LATE** es válida solo para los maestros del grupo de tratamiento que se inscribieron en el programa y que no se habrían inscrito si hubieran sido asignados al grupo de comparación. Los principios para estimar el **LATE** se aplican cuando hay incumplimiento en el grupo de tratamiento, en el de comparación, o en ambos al mismo tiempo. El **TOT** es simplemente un **LATE** en el caso más específico en que hay incumplimiento solamente en el grupo de tratamiento. **(Gertler et al., 2017).**

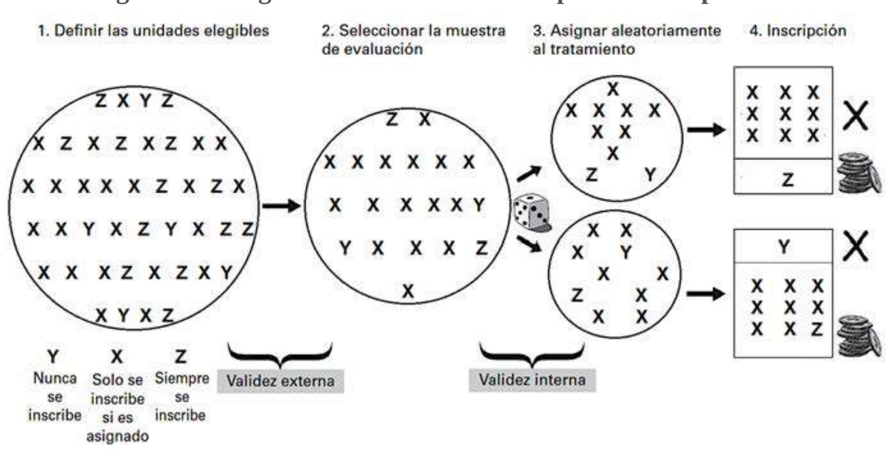
Asignación aleatoria de un programa y aceptación final

Suponga que se debe evaluar el impacto de un programa de capacitación en innovación para el empleo en los salarios de los individuos. El programa se asigna de forma *aleatoria* a nivel individual. El *grupo de tratamiento* recibe el programa, mientras que *el grupo de comparación* no lo recibe. Lo más probable es que se encuentren tres tipos de individuos en la población (**Gertler et al., 2017)**:

- *Inscritos si se lo ofrecen.* Son los individuos que cumplen con su asignación. Si se les asigna al grupo de tratamiento (*asignados al programa*), lo aceptarán y se inscribirán. Si se les asigna al grupo de comparación (*no asignados al programa*), no se inscriben.
- *Nuncas.* Son los individuos que *jamás se inscriben ni aceptan* el programa, aunque se les asigne al grupo de tratamiento. Si en efecto se les asigna a este último, serán incumplidores.
- *Siempres*. Estos individuos encontrar.n una manera de inscribirse en el programa o aceptarlo, aunque se les asigne al grupo de comparación. Si en efecto son asignados al grupo de comparación, serán incumplidores.

En el contexto de un programa de capacitación laboral, el grupo de los *Nuncas* puede estar formado por personas no motivadas que, aunque se les asigne un lugar en el curso, no se presentan. En cambio, los del grupo de los *Siempres* están tan motivados que encuentran una manera de entrar en el programa aunque originalmente se les haya asignado al grupo de comparación. El grupo de *Inscritos si se lo ofrecen* son los que se anotan en el curso si son asignados al, pero no buscan inscribirse si son asignados al grupo de comparación. **Ver Figura 6.1.**

Figura 6.1. Asignación aleatoria con cumplimiento imperfecto



Fuente: Gartler et al. (2017) con adaptación propia

La **Figura 6.1** presenta la asignación aleatoria del programa y de la inscripción final, o aceptación, cuando los tipos Inscritos si se lo ofrecen, *Nuncas* y *Siempres* están presentes. Supóngase que la población se compone de un **80%** de Inscritos si se lo ofrecen, **10%** de *Nuncas* y **10%** de *Siempres*. Si se toma una muestra aleatoria de la población para la muestra de evaluación, dicha muestra tendrá también alrededor de un **80%** de Inscritos si se lo ofrecen, **10%** de *Nuncas* y **10%** de *Siempres***.** Luego, si la muestra de evaluación se asigna de manera aleatoria a un grupo de tratamiento y a un grupo de comparación, una vez más debería haber cerca de un **80%** de Inscritos si se lo ofrecen, **10%** de *Nuncas* y **10%** de Siempres en ambos grupos. En el grupo asignado al tratamiento, se inscribirán los individuos Inscritos si se lo ofrecen y Siempres, y solo el grupo de *Nuncas* permanecerá al margen. En el grupo de comparación, los individuos de *Siempres* se inscribirán, mientras que los grupos de Inscritos si se lo ofrecen y *Nuncas* permanecerán fuera. Es importante recordar que si bien se sabe que en la población existen estos tres tipos de individuos, no es posible necesariamente distinguir el tipo de un individuo hasta que se observan ciertas conductas. En el *grupo de tratamiento*, se podrá identificar a los

tipos de Nuncas cuando no se inscriben, pero no se podrá distinguir entre los Inscritos si se lo ofrecen y los *Siempres*, dado que ambos tipos se inscribirán. En el *grupo de comparación*, se podrá identificar a los *Siempres* cuando se inscriben, pero no se podrá distinguir entre los Inscritos si se lo ofrecen y los *Nuncas*, dado que ninguno de los dos tipos se inscribirá.

Implicaciones de instrumentos débiles en los estimados

Un inconveniente del enfoque **IV**, es la dificultad potencial para encontrar un buen instrumento. Cuando el instrumento se correlaciona con las características no observadas que afectan el resultado, es decir, **cov (Z, ε) <> 0**, las estimaciones del efecto del programa estarán sesgadas. Además, si el instrumento solo se correlaciona débilmente con la variable de tratamiento ***T***, el estándar de la estimación **IV,** es probable que aumente el error porque el impacto predicho en el resultado, se medirá con menos precisión. La consistencia de la estimación **IV** (*sesgo asintótico*) es también probable que sea grande cuando ***Z*** y ***T*** están débilmente correlacionadas, incluso si la correlación entre ***Z*** y ε sea bajo. Este problema puede violar el supuesto subyacente a la estimación **IV** como se ve aquí. Como se mencionó en la sección anterior, asintóticamente, $\beta_{IV} = \beta + cov(Z, \varepsilon) / cov(Z, T)$**;** así, cuanto menor es el **cov (Z, T)**, mayor es el *sesgo asintótico* de β y más lejos del verdadero β**.**

Pruebas de los instrumentos débiles

No se puede comprobar si un instrumento específico satisface la restricción de exclusión; como se mencionó anteriormente, las justificaciones solo pueden hacerse a través de evidencia directa de cómo el programa y la participación evolucionaron. Con múltiples instrumentos, sin embargo, las pruebas cuantitativas existen y son también conocidas como pruebas de restricciones de sobreidentificación (*tests of overidentifying restrictions*) que implican los siguientes pasos **(Khandker et al., 2017)**:

1. Estimar la ecuación estructural de **2SLS** y obtenga los residuos ε_i'.
2. Realice la regresión de ε_i' (que incorpora toda la heterogeneidad no explicada por los instrumentos Z y otras variables exógenas X) en X y **Z**. Obtenga R^2.
3. Utilice la *hipótesis nula* de que todas las variables instrumentales no están *uncorrelacionadas* con los residuos, $n\,R^2 \sim \chi^2_q$ donde q es el número de variables instrumentales fuera del modelo menos el número total de variables explicativas endógenas. Si $n\,R^2$ es estadísticamente mayor que el valor crítico en un cierto nivel de significación **(5%)** en la que la distribución χ^2_q entonces se rechaza la *hipótesis nula*, y se puede concluir que al menos una de las variables instrumentales no es exógena.

Como se mencionó anteriormente, la estimación con el método **IV** del efecto del programa es en última instancia, una intención de tratar el cálculo de impacto, donde el efecto medido del programa, se aplicará solo a un subconjunto de participantes. La focalización imperfecta, es un caso donde solo los impactos **ITT** (*intent-to-treat*) pueden ser medidos; el investigador deberá buscar un *indicador exógeno* de participación que explique la *heterogeneidad no observada.* Un buen instrumento en este caso, debe satisfacer la restricción de exclusión y estar bien correlacionado con la participación. Sin embargo, es muy poco probable que el instrumento esté perfectamente correlacionado con la participación, por lo que solo un subconjunto de participantes sería recogido por el instrumento y el efecto del método **IV** resultante **(Khandker et al., 2017).**

Lo mismo ocurre, cuando se necesita un instrumento para corregir errores en la medición de la participación; resultaría en impactos similares de **ITT** relacionados con un subconjunto de participantes. El efecto resultante en el programa por el método **IV,** se aplicaría solo al subconjunto de participantes cuyo comportamiento se vería afectado por el instrumento. Una dificultad surge con la estimación estándar del método **IV** si los sujetos saben más sobre las retribuciones esperadas del programa, que el evaluador o investigador no pueda observar. Es decir, los sujetos anticipan retribu-

ciones del programa que el evaluador o investigador no puedo observar. En consecuencia, la *selección no observada* ocurre en la participación, debido a aquellos sujetos que terminan beneficiándose más del programa, dadas sus características ***X***, que también hace más probable que participe. Debido a que el instrumento ***Z*** afecta la participación, las *características no observadas* que impulsan la participación también se correlacionarán con ***Z***, y la estimación por el método **IV** estará sesgada. Sin embargo, la metodología como instrumento, causa cambios sistemáticos en las tasas de participación que se relacionan con ganancias anticipadas no observadas del programa. *Este cambio crea un sesgo al comparar participantes y no participantes.*

Imbens y Angrist (1994) abordan este problema introduciendo el concepto de promedio local efecto del tratamiento (**LATE.** *Local Average Treatment Effect*) En el caso especial, donde exista heterogeneidad en la respuesta de los sujetos para el programa, los métodos **IV** estiman consistentemente el efecto promedio del programa solo para aquellos cuya participación varía debido a cambios en el instrumento ***Z***. Específicamente, **LATE** estima el efecto del tratamiento solo para aquellos que deciden participar debido a un cambio en ***Z*** .

Una de las suposiciones subyacentes para **LATE** es la *monotonicidad*, o que un aumento en ***Z*** de ***Z*** = z a ***Z*** = z**’** lleva a algunos a participar pero a nadie a abandonar el programa. La participación ***T*** en este caso depende de ciertos valores de los instrumentos ***Z*** (digamos, ***Z*** = z vs. ***Z*** = z**’**), de modo que **P (*T = 1* | *Z = z*)** es la probabilidad de participar cuando ***Z*** = z’, y **P (*T = 1* | *Z* =** z**’)** es la probabilidad de participar cuando **Z = z′.** Como se discutió anteriormente, ***T*** es la variable de tratamiento igual a **1** para los participantes e igual a **0** para los no participantes. Los resultados ***Y*** y la participación ***T*** también son funciones de otras covariables observadas ***X***, que se han suprimido por simplicidad en la **ecuación 6.7.**

Tenga en cuenta que, recordando el capítulo 4, **P (*T = 1* | *Z* =** z**)** y **P (*T = 1* | *Z* =** z′**)** también pueden interpretarse como la propensión de puntajes de participación basados en los instrumentos ***Z***, es decir, **P (**z**)** y **P (**z**’),** respectivamente, así:

$$\beta_{IV,\,LATE} = \frac{E(Y|P(Z)=P(z)) - E(Y|P(Z)=P(z'))}{P(z)-P(z')} \quad \ldots\ldots\ldots(6.7)$$

El denominador en la **ecuación 6.7**, es la diferencia en la probabilidad de participar en el programa (probabilidad de $T=1$) bajo los diferentes valores de los instrumentos Z = z a Z = z**'.** Usando la **ecuación 6.7**, se puede estimar el **LATE** utilizando métodos lineales **IV**, por lo que:

- En la primera etapa, la participación en el programa T se estima en función de los instrumentos Z para obtener el puntaje de propensión, **P' (Z) = P' (T = 1 | Z).**
- En segundo lugar, se puede estimar una regresión lineal del resultado $Y_{it} = [\, T_i \cdot Y_i(1) + 1 - T_i\,) \cdot Y_i(0)\,]$ sobre **P' (Z).** La interpretación del efecto estimado del programa β_{IV}' es el cambio promedio en los resultados Y de un cambio en la participación de la propensión de puntaje **P'(*Z*)** manteniendo otras covariables observadas X fijas.

La estimación de impacto bajo asignación aleatoria con cumplimiento imperfecto

Después de establecer la diferencia entre asignar un programa y la inscripción o aceptación en la práctica, se estimará el **LATE** del programa. Esta estimación se lleva a cabo en dos pasos. (**Ver Figura 6.2**):

Figura 6.2. Estimación del efecto local promedio del tratamiento bajo asignación aleatoria con cumplimiento imperfecto

	Grupo asignado al tratamiento	Grupo no asignado al tratamiento	Impacto
	Porcentaje inscrito= 90% Media (*Y*) para los asignados a tratamiento = 110	Porcentaje inscrito = 10% Media (*Y*) para los no asignados a tratamiento = 70	Δ porcentaje de inscritos = 80% ΔY = ITT = 40 LATE = 40%/80% = 50
Nunca se inscribe	Y	Y	-------
Solo se inscribe si es asignado	XXXXXX XXXXXX	XXXXXX XXXXXX	XXXXXX XXXXXX
Siempre se inscribe	Z	Z	-------

Nota: La estimación de la intención de tratar (ITT) se obtiene comparando los resultados de los individuos asignados al grupo de tratamiento con los de aquellos asignados al grupo de comparación, independientemente de la inscripción en la práctica. La estimación del efecto local promedio del tratamiento (LATE) es el impacto del programa en los que se inscriben solo si son asignados al programa (Inscritos si se lo ofrecen). La estimación LATE no proporciona el impacto del programa en aquellos que nunca se inscriben (Nuncas) o en aquellos que siempre se inscriben (Siempres).
Δ = impacto causal; Y = resultado
Las figuras que aparecen con el fondo sombreado corresponden a los que se inscriben.
Fuente: Gertler et al. (2017) con adaptación propia.

a. Para estimar los impactos del programa bajo la asignación aleatoria con cumplimiento imperfecto, primero se estima el impacto de la **ITT**. Se debe recordar que se trata solo de la diferencia en el indicador de resultados **(*Y*).** Por ejemplo, si el salario medio **(*Y*)** del grupo de tratamiento es **110** y el salario medio del grupo de comparación es **70**, la estimación de la **ITT** del impacto sería de **40 (110- 70).**

b. En segundo lugar, habría que recuperar la estimación del **LATE** para el grupo de Inscritos si se lo ofrecen de la estimación **ITT.** Para esto, se debe identificar de dónde proviene la diferencia de **40.** Se procede por eliminación:

- Primero, se sabe que la diferencia no puede ser causada por diferencias entre los sujetos que nunca se inscriben (*los Nuncas*) en los *grupos de tratamiento* y de comparación. Esto se debe a que *los*

Nuncas jamás se inscriben en el programa, de modo que para ellos es igual estar en el grupo de tratamiento que en el de comparación.

- En segundo lugar, se sabe que la diferencia no puede ser producida **(40)** por diferencias entre los sujetos de *Siempres* en los *grupos de tratamiento y de comparación* porque estos siempre se anotan en el programa. Para ellos tampoco hay diferencia entre estar en el grupo de tratamiento o el grupo de comparación.

c. Por lo tanto, la diferencia en los resultados entre ambos grupos debe necesariamente provenir del efecto del programa en el único grupo afectado por su asignación al grupo de tratamiento o de comparación, es decir, el grupo de Inscritos si se lo ofrecen . Si se puede identificar *a los Inscritos* si se lo ofrecen en ambos grupos, será fácil estimar el impacto del programa en ellos. En realidad, aunque se sabe que estos tres tipos de individuos existen en la población, no se puede separar a los individuos en función de si son *Inscritos si se lo ofrecen*, *Nuncas o Siempres*.

En el grupo que fue asignado al tratamiento, se puede identificar a los Nuncas (porque no se han inscrito), pero no se puede diferenciar entre los Siempres y los Inscritos si se lo ofrecen (porque ambos están inscritos). En el grupo de comparación, se puede identificar el *grupo de Siempres* (porque se inscriben en el programa), pero no es *posible diferenciar entre los Nuncas y los Inscritos si se lo ofrecen.* Sin embargo, una vez que se observa que el **90%** de las unidades en el grupo asignado al tratamiento en efecto se inscribe, se puede deducir que el **10%** de las unidades de nuestra población debe estar formada *Nuncas* (es decir, el porcentaje de individuos del grupo asignados al tratamiento que no se inscribieron). Además, si se observa que el **10%** de las unidades del grupo de comparación se inscribe, se sabe que el **10%** son *Siempres* (una vez más, el porcentaje de individuos de nuestro grupo que no fue asignado al tratamiento y que sí se inscribió). Esto deja al **80%** de las unidades en el grupo de Inscritos si se lo ofrecen. Se sabe que el impacto de **40** proviene de una diferencia en la inscripción en el **80%** de las unidades de la muestra que corresponde a

Inscritos si se lo ofrecen. Si el **80%** de las unidades es responsable de un impacto promedio de **40** en el conjunto del grupo asignado al tratamiento, el impacto en ese **80%** de Inscritos si se lo ofrecen debe ser **40/0.8**, o **50.** Dicho de otra manera, el impacto del programa para los Inscritos si se lo ofrecen es de **50**, pero cuando este impacto se distribuye en el conjunto del grupo asignado al tratamiento, el efecto promedio se diluye debido al **20%** que no cumplió con la asignación aleatoria original.

Recuerde que uno de los problemas básicos de la autoselección en los programas es que no siempre se puede saber por qué algunas personas deciden participar y otras no. Cuando se lleva a cabo una evaluación donde las unidades están asignadas de forma aleatoria, pero la participación en la práctica es voluntaria o existe una forma en que las unidades del grupo de comparación participen en el programa, se presenta un problema similar, a saber, que no siempre se entenderá la conducta que determina si un individuo se comporta como un *Nunca, un Siempre, o un Inscrito si se lo ofrecen*.

Sin embargo, si la falta de cumplimiento no es demasiado severo, la asignación aleatoria sigue proporcionando un instrumento útil para la evaluación del impacto. El aspecto negativo de la asignación aleatoria con cumplimiento imperfecto es que esta estimación de impacto ya no es válida para el conjunto de la población. En cambio, la estimación debería interpretarse como una estimación local que se aplica solo a un subgrupo específico dentro de la población designada, los Inscritos si se lo ofrecen.

La *asignación aleatoria* de un programa tiene dos características importantes que permiten estimar el impacto cuando hay *cumplimiento imperfecto*:

1. Puede servir para predecir la inscripción en el programa si la mayoría de las personas se comportan como Inscritos si se lo ofrecen, y se inscriben en el programa cuando se les asigna al tratamiento y no se inscriben cuando no se les asigna.

2. Dado que los dos grupos (asignados y no asignados al tratamiento) se generan mediante un proceso aleatorio, las características de los indi-

viduos en los dos grupos no están correlacionadas con ningún otro factor, como la habilidad o la motivación, que pueda influir también enlos resultados **(*Y*).**

En términos estadísticos, la asignación aleatoria sirve como **IV**. Se trata de una variable que predice la inscripción real de unidades en un programa, pero que no está relacionada con otras características de los individuos que puedan estar vinculadas a los resultados. Aunque en parte la decisión de los individuos de inscribirse en un programa no puede estar controlada por los administradores del programa, otra parte de la decisión sí está bajo su control. Concretamente, la parte de la decisión que puede controlarse es la asignación a los grupos de tratamiento y comparación. En la medida en que la asignación a los grupos de tratamiento y de comparación predice la inscripción final en el programa, la asignación aleatoria se puede usar como un instrumento para predecir la inscripción final. Tener esta **IV** permite recuperar las estimaciones del **LATE** de las estimaciones del efecto de **ITT** para el tipo de unidades Inscritos si se lo ofrecen.

Una variable **IV** debe satisfacer dos condiciones básicas **(Gertler et al., 2017):**

1. No debería estar correlacionada con las características de los *grupos de tratamiento y de comparación.* Esto se consigue asignando el tratamiento aleatoriamente a las unidades en la muestra de evaluación. Esto se conoce como exogeneidad. Es importante que la VI no influya directamente en el resultado de interés. Los impactos deben ser causados únicamente a través del programa que nos interesa evaluar.
2. Debe influir en las tasas de participación de los grupos de tratamiento y comparación de manera diferente. Normalmente se piensa en aumentar la participación en el grupo de tratamiento. Esto se puede verificar constatando que la participación es más alta en el grupo de tratamiento que en el de comparación. Esta condición se conoce como *relevancia.*

Evaluación de la IV

Se requiere asumir que al menos uno de los instrumentos es exógeno y luego evaluar las restricciones de *sobreidentificación*. La prueba se denomina de esta manera porque para obtener el estimador **IV** se requiere al menos un instrumento. Si se dispone de más instrumentos, entonces el sistema tiene más ecuaciones que incógnitas, por lo cual estaría *sobreidentificado*. Cabe destacar que la prueba de sobreidentificación tan solo es válida bajo el supuesto de que los efectos son homogéneos. Si los efectos son heterogéneos, el grupo de cooperativos será distinto para cada instrumento lo que implica que el efecto local también lo será. En la **Tabla 6.1 y 6.2** se presentan pruebas de relevancia y exogeneidad adicionales, que se pueden utilizar para evaluar la *validez* de una **IV.**

Tabla.6.1. Pruebas de relevancia

Nombre de la prueba	Estadístico utilizado	Descripción	Pasos	
1. Regresión simple	LF	El objetivo de esta prueba es simplemente evaluar si existe una relación estadísticamente significativa entre los instrumentos *(Z)* y la variable endógena *(D)*. Esto se hace corriendo una regresión donde la variable endógena es la variable dependiente y los regresores son los instrumentos y las otras variables explicativas del modelo *(X)* diferentes al indicador de tratamiento. Si se encuentra que los coeficientes asociados a los instrumentos y el modelo conjuntamente son significativos, hay evidencia empírica para garantizar la relevancia de los instrumentos.	1. Suponga que nos interesa evaluar el impacto del programa *(D)* sobre la variable de interés *(Y)*. Además del programa, los otros determinantes de la variable *(Y)* están incluidos en *(X)*.	$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i + u_i$
			2. Se estima una regresión donde la variable independiente es la participación en el programa *(D)* y las variables independientes son los instrumentos *(Z)* y los demás regresores *(X)*.	$D_i = \lambda_0 + \lambda_1 Z_i + \lambda_2 X_i + \varepsilon_i$
			3. Se evalúa la significancia estadística de los coeficientes asociados a los instrumentos.	$H_0 = \lambda_1 = 0$ $H_a = \lambda_1 \neq 0$
2. Pruebas de identificación (Cragg-Donald y prueba canónica de Anderson)	χ^2	Las pruebas de relevancia se pueden dividir en dos. Pruebas de identificación y pruebas de instrumentos débiles. Las primeras tienen como objetivo identificar si existe una relación estadísticamente significativa entre los instrumentos y la variable endógena. Dentro del grupo de pruebas de identificación están la prueba Cragg-Donald y la prueba canónica de Anderson.	1. Se identifica la ecuación que se quiere estimar y se seleccionan los posibles instrumentos.	$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i + u_i$
				Posibles instrumentos: $Z_1, Z_2, \ldots, Z_n$
			2. Se hace una estimación de la primera etapa. La variable independiente es la endógena y los regresores son los instrumentos y las otras variables explicativas X_i, incluidos en Φ_i.	$D_i = \Pi\Phi_i + \varepsilon_i$

Nombre de la prueba	Estadístico utilizado	Descripción	Pasos	
2. Pruebas de identificación (Cragg-Donald y prueba canónica de Anderson)	χ^2	Estas pruebas tienen por objetivo identificar si la matriz de coeficientes de la primera etapa tiene rango completo. Si ese es el caso, los instrumentos son relevantes.	3. Se verifica que la matriz Π tenga rango completo. Para esto, la prueba canónica de Anderson y la prueba Cragg-Donald construyen un estadístico de prueba distinto pero en esencia las dos pruebas son bastante parecidas.	$H_0 = \Pi$ tiene rango completo $H_a = \Pi$ no tiene rango completo
3. Prueba de instrumentos débiles (Prueba de Stock y Yogo)	χ^2	Aun cuando existe una relación estadísticamente significativa entre los instrumentos y la variable endógena, puede haber problemas. No basta con que exista una relación significativa entre los instrumentos y la variable endógena, se necesita que esta relación sea fuerte. Stock y Yogo desarrollaron esta prueba considerando que un conjunto de instrumentos es débil si cumple dos características. Nosotros utilizaremos únicamente la primera, que se refiere a un instrumento como débil si la diferencia entre el estimador de MC2E y los coeficientes verdaderos sobrepasa un límite determinado.	1. Se identifica la ecuación de interés y se estiman sus parámetros por MCO.	$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i + u_i \rightarrow \hat{\beta}_i^{MCO}$
			2. Se estima la ecuación de interés por MC2E.	$D_i = \lambda_0 + \lambda_1 Z_i + \lambda_2 X_i + \varepsilon_i \rightarrow \hat{D}_i$ *por MCO* $Y_i = \beta_0 + \beta_1 X_i + \beta_2 \hat{D}_i + u_i \rightarrow \hat{\beta}_i^{MC2E}$
			3. Se construye el estadístico de prueba que depende del número de instrumentos *(k)*, el número de variables endógenas *(n)*, el tamaño máximo de sesgo deseado *(r)* y la relación entre el estimador de MCO y el de MC2E.	$SY = SY\left(r, n, k, \frac{\beta^{MCO}}{\beta^{MC2E}}\right)$
			4. Se compara este valor con el valor crítico de la prueba Cragg-Donald. De esta manera, si se rechaza la hipótesis nula, podemos garantizar que los instrumentos son lo suficientemente fuertes como para no generar un sesgo mayor a *r*.	H_0: Instrumentos Z_i débiles (*SY*> valor crítico *Cragg-Donald*) H_a: Instrumentos Z_i no débiles (*SY*< valor crítico *Cragg-Donald*)

Fuente: Bernal y Peña (2011).

Tabla. 6.2. Pruebas de exogeneidad

Nombre de la prueba	Estadístico utilizado	Descripción	Pasos	
1. Prueba de Sargan	χ-cuadrado	Esta prueba evalúa las restricciones de sobreidentificación (en el caso en que se dispone de más instrumentos que variables explicativas endógenas). Se basa en la observación de que los residuales no deberían estar correlacionados con los instrumentos si éstos son realmente exógenos.	Se hace la estimación de la ecuación de interés por MC2E en el modelo *exactamente* identificado.	$Y_i = \beta_0 + \beta_1 D_i + \beta_2 X_i + u_i \rightarrow \hat{\beta}_i^{MC2E}$
			Se estima una regresión de los residuales de la regresión anterior sobre todas las variables exógenas. Z_i es un vector de instrumentos para el individuo *i*.	$\hat{u}_i = \lambda_0 + \lambda_1 Z_i + \lambda_2 X_i + \varepsilon_i$
			Se construye el estadístico de Sargan utilizando el número de observaciones *(n)*, y el R^2 de la regresión en el paso 2. Los grados de libertad corresponden al número de restricciones de sobreidentificación, es decir, el número de instrumentos, menos el número de variables endógenas.	$S = n * R^2 - X^2_{L-K}$ H_0: Instrumentos no correlacionados con el error H_a: Instrumentos correlacionados con el error Si se rechaza la H_0, implicaría que al menos uno de los instrumentos no es exógeno.

Nombre de la prueba	Estadístico utilizado	Descripción	Pasos	
2. Prueba J (Hansen)	χ-cuadrado	La prueba de sobreidentificación J de Hansen es una extensión de la prueba propuesta por Sargan para el caso de estimación lineal por variables instrumentales. En particular, la hipótesis nula de las dos pruebas es idéntica pero el estimador de variables instrumentales se estima en este caso por el método generalizado de momentos (GMM). Esta prueba es robusta a presencia de heterocedasticidad. La estimación de GMM se basa en el momento poblacional dictado por la disponibilidad de un instrumento válido: $E(Z'u) = 0$.	Se estima la ecuación de interés mediante variables instrumentales utilizando el método generalizado de momentos (GMM). La estimación por GMM minimiza el análogo muestral del momento poblacional dictado por variables instrumentales, es decir, $E(Z'u) = 0$. Como el modelo está sobreidentificado (hay más momentos que parámetros) entonces GMM minimiza: $J = u'ZW^{-1}Z'u \rightarrow \hat{\beta}_1^{GMM}$ Donde Z es la matriz de instrumentos, u es el vector de errores y W es la matriz de varianza. Por tanto, la función objetivo pesa más los momentos estimados con mayor precisión y viceversa.	
			Se construye el estadístico de Hansen que corresponde al número de observaciones n por el valor minimizado de la función objetivo de GMM. Este estadístico está distribuido como una ji-cuadrado, con grados de libertad L-K, donde L es el número de instrumentos y K es el número de variables endógenas.	$\chi^2_{L-K} = nJ(\beta_{GMM})$

Fuente: Bernal y Peña (2011).

Otra manera de evaluar la estrategia de identificación propuesta a través de la **IV** es aplicar una *prueba de falsificación* como aquellas discutidas previamente. En este caso, por ejemplo, se podría aplicar una prueba de pseudoresultada en particular, podría estimarse el modelo por variables instrumentales utilizando como variable de resultado una variable que estamos seguros de que no pudo estar afectada por el programa.

En el caso del programa de estímulos a la innovación (**pei**), se espera que el programa tenga efecto sobre el los ingresos en las **pbt** después de la participación en el programa, pero no debería tener efecto sobre la **pbt** antes. En este caso, podríamos estar preocupados por que los instrumentos si estén correlacionados con cosas como la importancia que los CEOS atribuyen al desarrollo de la **pbt**, con la habilidad innata de liderazgo (que no tenemos registrada en la base de datos) o con otras cualidades que también determinan el desarrollo de la **pbt**. Podríamos utilizar los ingresos por innovación en el primer año de operación como prueba de validez, porque ese primer año debería estar influenciado por las características que si observamos del CEO y sus recursos, pero no por las inversiones posteriores de la **pbt** como la participación en **pei**.

Interpretación de la estimación del efecto promedio del tratamiento local

La diferencia entre la estimación de un ATE y la estimación de un **LATE** es especialmente importante cuando se trata de interpretar los resultados de una evaluación **(Gertler, et al., 2017):**

- En primer lugar, debe reconocerse que los individuos que cumplen en un programa (el tipo Inscritos si se lo ofrecen) son diferentes de los individuos que no cumplen (*los tipos Nuncas y Siempres*). Concretamente, en el grupo de tratamiento, los no cumplidores/no participantes (*Nuncas*) pueden ser aquellos que esperan ganar poco con la intervención. En el *grupo de comparación,* los no cumplidores/participantes (*Siempres*) probablemente constituyan el grupo de individuos que esperan el mayor beneficio de participar.
- En segundo lugar, se sabe que la estimación **LATE** proporciona el impacto para un subgrupo particular de la población: tiene en cuenta solo al subgrupo que no se ve afectado por ningún tipo de incumplimiento. En otras palabras, tiene en cuenta solo el tipo Inscritos si se lo ofrecen. Dado que el tipo Inscritos si se lo ofrecen es diferente de los *Nuncas* y de los *Siempres*, el impacto que se halla a través de la estimación **LATE** no se aplica a los tipos *Nuncas o Siempres*.

La estimación del **LATE** se aplica únicamente a un subconjunto específico de la población, a saber, aquellos tipos que no están afectados por la falta de cumplimiento, es decir, solo el *tipo cumplidor* y no debería extrapolarse a otros subconjuntos de la población.

Efecto de tratamiento marginal (*Marginal Treatment Effect*)

El efecto de tratamiento marginal (**MTE.** *Marginal Treatment Effect*)**,** introducido en el capítulo 3, es la forma límite del **LATE** discutido por **Heckman y Vytlacil (2005)** y **Todd (2007)** como un método para estimar los efectos del tratamiento cuando la *exogeneidad condicional* no se cumple. Como se mencionó anteriormente, el **MTE** es la ganancia promedio en resultados para los participantes cercanos al umbral o en el margen de participación, dado un conjunto de características observadas y condicionamiento sobre un conjunto de características no observadas en la ecuación de participación. Siguiendo a **Heckman y Vytlacil (2000),** el **MTE** se puede escribir como:

$$\text{MTE} = E(\, Y_1 (1) - Y_1 (0) \mid X_1 = x,\, U_1 = u) \ldots\ldots\ldots\ldots\ldots (6.8)$$

En la **ecuación 6.8** $\boldsymbol{Y_1}$ **(1)** es el resultado para todos los sujetos en tratamiento, $\boldsymbol{Y_i}$ **(0)** es el resultado para todos los que no lo reciben, $\boldsymbol{X_i = x}$ son características observadas para el sujeto $\boldsymbol{i}$, y $\boldsymbol{U_i = u}$**,** $\boldsymbol{U_i} \in$ **(0,1**) son características no observadas, para el sujeto $\boldsymbol{i}$ que también determinan participación. Observando el efecto $\boldsymbol{U_1}$ de en la participación $\boldsymbol{T_i}$ (ver capítulos anteriores que $\boldsymbol{T_i = 1}$ para participantes y $\boldsymbol{T_i = 0}$ para no participantes), **Heckman y Vytlacil (2000)** suponen que $\boldsymbol{T_i}$ es generado por una variable latente $\boldsymbol{T_i^*}$**.** Esta ecuación también se conoce como un *modelo de índice latente lineal* (*linear latent index model*, ver **Heckman y Robb, 1985; Imbens y Angrist, 1994**) Así:

$$\boldsymbol{T_i^* = \mu_t (Z_i) - U_i}$$
$$T_i = 1 \text{ if } T_i^* > 0\ ,\ T_i = 0 \text{ if } T_i^* <= 0 \ldots\ldots\ldots\ldots\ldots (6.9)$$

Donde $\boldsymbol{Z_i}$ son instrumentos observados que afectan la participación y $\boldsymbol{\mu_t (Z_i)}$ es una función que determina los posibles resultados $\boldsymbol{Y}$ de $\boldsymbol{Z}$ que están condicionados a la participación. Sujetos con *características no*

observadas o cercanas a cero, por lo tanto, son las más probables para participar en el programa (T_i más cerca de **1**) y sujetos con ***u*** cerca de **1** como los menos propensos a participar. El **MTE** para individuos con $U_i = u$ cerca de **0** por lo tanto, refleja el efecto del tratamiento promedio (**ATE**) para las personas más inclinadas a participar, y el **MTE** para individuos con $U_i = u$ cercano a **1** representa el **ATE** para sujetos con menos probabilidades de participar.

¿Por qué es útil el **MTE** para comprender los efectos del tratamiento? Además, si tanto el **MTE** y el **LATE** examinan el impacto variable de las características no observadas en la participación, ¿cuál es la diferencia entre ellos? Tanto el **MTE** como el **LATE** permiten individuos para anticipar su retribución en ***Y*** sobre la base de las características no observadas. Sin embargo, justo como el **LATE** es una versión más fina del efecto del tratamiento sobre el tratado (**TOT**. *Treatment Effect On Treatmet*) **(Heckman et al.; 1997),** el **MTE** es la forma límite del **LATE** y define mucho mejor el efecto del tratamiento, más preciso que el **LATE** para un cambio infinitesimal en **Z (Heckman y Vytlacil, 2000).**

Una propiedad útil del **MTE (Heckman y Vytlacil, 2000, 2005)** es que el **ATE**, **TOT** y **LATE** pueden obtenerse al integrarse en diferentes regiones del **MTE**. El **ATE,** que, como se discutió en el capítulo 3, es el efecto promedio para toda la población,es decir, el efecto del programa para un sujeto extraído al azar de la población, se puede obtener integrando el **MTE** sobre todo el soporte ($u = 0$ **a** $u = 1$).

El **TOT**, que es el *efecto de tratamiento promedio* para aquellos que eligen participar, se puede obtener integrando **MTE** de $u = 0$ a $u = P(z)$. Como se describió anteriormente, $P(z)$ es la propensión de puntaje, o probabilidad, de participar cuando el instrumento $Z = z$. Por lo tanto, el **TOT** es el efecto del tratamiento para individuos cuyas características no observadas hacen que sean más propensos a participar en el programa.

Finalmente, si se supone (como anteriormente) que el instrumento ***Z*** puede tomar valores $Z = z'$ y $Z = z$, y también que $P(z') < P(z)$, entonces **LATE** integra a **MTE** de $u = P(z')$ a $u = P(z)$. Este resultado ocurre, porque, cuando $P(z') < P(z)$, algunos sujetos, quienes no habían partici-

pado cuando ***Z* = *z*'**, participarán cuando ***Z*** = *z*, pero ningún sujeto que participó en **Z = z′** abandonará el programa cuando ***Z*** = *z*.

Siendo así entonces ¿cómo estimar el **MTE**? **Heckman y Vytlacil (2000)** proponen dos etapas del estimador de **IV** local:

$$\beta_{LIV,MTE} = \lim_{P(z')\to P(z)} \frac{E(Y|P(Z)=P(z)) - E(Y|P(Z)=P(z'))}{P(z)-P(z')} \quad \text{……….(6.10)}$$

El enfoque es similar a la estimación de **LATE** discutido anteriormente:

1. En la primera etapa, la participación en el programa todavía se estima en función de los instrumentos ***Z***, para obtener la propensión de puntaje **P' (*Z*)**.
2. En la segunda etapa, sin embargo, una regresión lineal no paramétrica se debe estimar del resultado $\boldsymbol{Y_i = [\, T_i \cdot Y_i\,(1) + (1 - T_i) \cdot Y_i\,(0)]}$ en **P' (*Z*)**. La evaluación de esta función a diferentes valores de la propensión de puntaje produce la función **MTE**.

El método **IV** local es diferente del método **IV** utilizado para estimar **LATE**, en el sentido de que el método **IV** local estima el cambio promedio en ***Y*** alrededor de un vecindario local de **P (Z)**, mientras que **LATE** se estima globalmente sobre el soporte (esta diferencia puede verse también comparando las **ecuaciones 6.7** y **6.10**).

Los enfoques para estimar **MTE** son nuevos y evolucionan. **Moffitt (2008)** también propone un método no paramétrico para estimar el **MTE**. En lugar de un procedimiento de dos pasos, donde la participación se instrumenta primero y posteriormente, el cambio promedio se calcula en ***Y*** sobre la base de la participación prevista, **Moffitt (2008)** estima el resultado y las ecuaciones de la participación, de forma conjunta a través de *mínimos cuadrados no lineales*. Este método, relaja algunos de los supuestos integrados en los modelos de *índice lineal* **IV** y *latente*. Sin embargo, existen muy pocas aplicaciones de **MTE** hasta ahora, particularmente en los países en desarrollo **(Khandker et al., 2017).**

Promoción aleatoria como variable instrumental

Ya se ha expuesto cómo estimar el impacto sobre la base de la asignación aleatoria del tratamiento, aun cuando el cumplimiento con los grupos de tratamiento y comparación originalmente asignados sea imperfecto. A continuación, se propone un enfoque muy similar que se puede aplicar a la evaluación de programas que tienen *elegibilidad universal o inscripción abierta,* o en los que el administrador del programa no puede controlar quién participa y quién no. Este enfoque, *denominado promoción aleatori*a, proporciona un estímulo más para que un conjunto aleatorio de unidades se inscriba en el programa. Esta promoción aleatoria sirve como **IV**. Sirve como una fuente externa de variación que afecta la probabilidad de recibir tratamiento, pero no está relacionada de ninguna forma con las características de los participantes.

Los programas de participación voluntaria suelen permitir que los individuos que se interesan en el programa decidan por sí mismos si quieren inscribirse y participar. Considere una vez más, en el programa de formación capacitación de la innovación tratado anteriormente, aunque esta vez la asignación aleatoria no es posible y cualquier individuo que desee inscribirse en el programa puede hacerlo. De manera muy parecida a la del ejemplo, se prevé encontrar diferentes tipos de personas: *cumplidores, un grupo de Siempres y un grupo de Nuncas*. Así**, (Gertler et al., 2017):**

- *Siempres.* Los individuos que siempre se inscribirán en el programa.
- *Nuncas.* Los individuos que jamás se inscribirán.
- *Cumplidores* o Inscritos si se promueve. En este contexto cualquierisujeto que quiera inscribirse en el programa puede hacerlo. Sin embargo, algunos sujetos pueden estar interesados en inscribirse, pero por diversos motivos no tienen suficiente información o el incentivo correcto para hacerlo. En este caso, los *cumplidores* son aquellos que se inscriben si se promueve. Se trata de un grupo de sujetos que se anotan en el programa solo si se les ofrece un incentivo adicional, un estímulo o motivación que los impulse a participar. Sin

este estímulo adicional, los *Inscritos* si se promueve sencillamente quedarían fuera del programa.

Para volver al ejemplo de la formación en la capacitación de la innovación, si la agencia que la organiza está bien financiada y tiene suficiente capacidad, es posible que despliegue una política de *puertas abiertas* y trate a todas las personas desempleadas que quieran participar. Sin embargo, es poco probable que todas las personas desempleadas quieran participar o incluso que sepan que el programa existe. Es posible que, algunas personas desempleadas tengan reparos para inscribirse porque saben muy poco acerca del contenido de la formación y les cuesta obtener información adicional.

Suponga que la agencia de capacitación en innovación, contrata a un trabajador de extensión comunitaria para que se pasee por la ciudad a fin de alentar a un grupo de personas desempleadas seleccionado de forma aleatoria para que se inscriban en el programa de formación laboral. Con la lista de personas desempleadas elegidas de manera aleatoria, llama a sus puertas, describe el programa de formación y les ofrece ayuda para inscribirse en ese mismo momento. La visita es una forma de promoción o estímulo para participar en el programa. Desde luego, no se puede obligar a nadie a participar. Además, las personas desempleadas que el trabajador de extensión comunitaria no visita también pueden inscribirse, aunque tendrán que ir personalmente a la agencia para hacerlo. Por lo tanto, ahora hay *dos grupos de personas desempleadas:* aquellas que fueron asignadas de modo aleatorio a una visita del trabajador comunitario y aquellas que aleatoriamente no fueron visitadas. Si el esfuerzo de extensión es efectivo, la tasa de inscripción entre las personas desempleadas que fueron visitadas debería ser superior a la tasa entre las personas desempleadas que no fueron visitadas.

Suponga ahora, en cómo se puede evaluar este programa de capacitación. No se puede simplemente comparar a las personas desempleadas que se inscriben con aquellas que no se inscriben. Esto se debe a que los desempleados que se inscriben probablemente sean muy diferentes de aquellos que no lo hacen, tanto en sus características observables como no observa-

bles. Es posible que tengan un nivel educativo mayor o menor (esto puede observarse con facilidad) y probablemente estén más motivados y deseosos de encontrar un empleo (esto es difícil de observar y medir). Sin embargo, hay una variación adicional que se puede explotar para encontrar un grupo de comparación válido. Piénsese si se puede comparar el grupo de sujetos que fueron asignados aleatoriamente para recibir una visita del trabajador de extensión con el grupo que no fue visitado. Dado que los grupos con *promoción y sin promoción* fueron determinados de forma *aleatoria*, ambos contienen composiciones idénticas de personas muy motivadas (Siempres) que se inscribirán independientemente de que el trabajador de extensión llame a su puerta o no.

Ambos grupos también contienen *personas no motivadas* (*Nuncas*) que no se inscribirán en el programa, a pesar de los esfuerzos del trabajador de extensión. Por último, si el trabajador de extensión es efectivo motivando a las personas a inscribirse, algunos (*Inscritos si se promueve*) se anotarán en el programa si el trabajador de extensión los visita, pero no lo harán si no reciben dicha visita. Dado que el trabajador de extensión visitó a un grupo de individuos asignados de manera aleatoria, puede derivarse una estimación **LATE,** como se señalaba anteriormente. *La única diferencia es que en lugar de asignar el programa de modo aleatorio, se lo está promoviendo aleatoriamente*. Siempre que los Inscritos si se promueve (que se inscriben cuando se hace contacto con ellos pero no se anotan si no hay contacto) sean lo suficientemente numerosos, entre el grupo con la *promoción y el grupo sin la promoción* habrá variaciones que permitirán identificar el impacto de la formación en los *Inscritos* si se promueve. En lugar de cumplir la asignación del tratamiento, los *Inscritos* si se promueve ahora cumplen con la promoción.

Para que esta estrategia funcione, la actividad de promoción tiene que ser efectiva y aumentar la inscripción considerablemente en el grupo de Inscritos si se promueve. Al mismo tiempo, las actividades de promoción en sí mismas no deberían influir en los resultados finales de interés (como los ingresos), dado que al final lo que interesa sobre todo es estimar el

impacto del programa de formación y no el impacto de la estrategia de promoción en los resultados finales.

Por ejemplo, si el trabajador de extensión ofreció grandes cantidades de dinero o favores especiales a los **pbt** para conseguir que se inscribieran, sería difícil saber si algún cambio posterior en los ingresos fue causado por la formación o por la actividad de promoción.

La *promoción aleatoria* es una estrategia creativa que genera el equivalente de un grupo de comparación para los fines de la evaluación de impacto. Se puede usar cuando un programa tiene inscripción abierta y es posible organizar una campaña de promoción destinada a una muestra aleatoria de la población de interés. La promoción aleatoria es otro ejemplo de VI que permite evaluar el impacto de manera no sesgada. Sin embargo, una vez más, como sucede con la *asignación aleatoria con cumplimiento imperfecto*, las evaluaciones de impacto que dependen de la promoción aleatoria proporcionan una estimación **LATE:** una estimación local del efecto en un subgrupo específico de la población, el grupo de Inscritos si se promueve. Como sucedió antes, esta estimación **LATE** no puede extrapolarse directamente al conjunto de la población, dado que los grupos de *Siempres y Nuncas* probablemente sean bastante diferentes del grupo de Inscritos si se promueve (**Gertler et al., 2017).**

Lo que representan las promociones en el IV

La promoción aleatoria pretende aumentar la aceptación de un programa voluntario en una submuestra de la población seleccionada aleatoriamente. La promoción puede adoptar diversas formas. Por ejemplo, puede que se decida iniciar una campaña de información para llegar a aquellas personas que no se han inscrito porque no lo sabían o porque no entienden cabalmente el contenido del programa. También, se pueden ofrecer incentivos para inscribirse, como pequeños obsequios o premios, o facilitando el transporte. Como se señaló de manera más general en el caso de las **IV**, para que el método de promoción aleatoria genere una estimación válida del

impacto del programa debe cumplirse una serie de condiciones **(Gertler et al., 2017):**

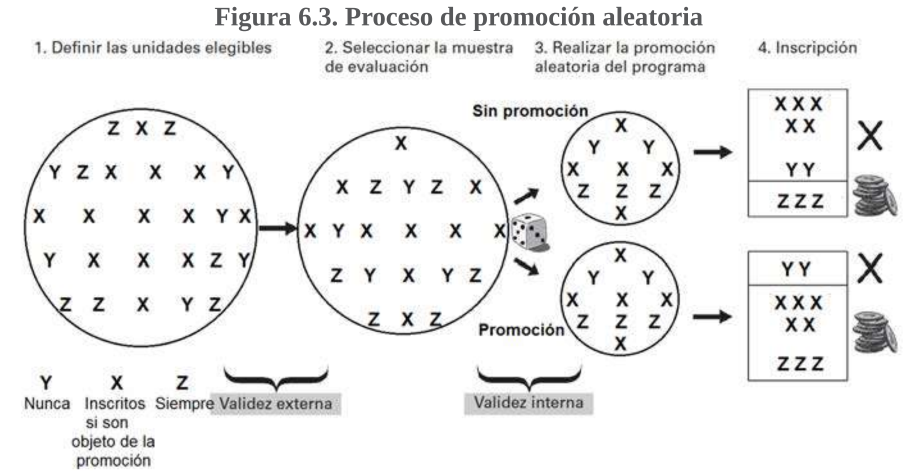
1. Los grupos que son objeto y no objeto de la promoción deben ser similares. Es decir, las características promedio de los dos grupos deben ser estadísticamente equivalentes. Esto se consigue asignando de forma aleatoria las actividades de extensión o promoción entre las unidades de la muestra de evaluación.
2. La propia promoción no debería influir directamente en los resultados de interés. Este es un requisito crítico, de modo que se pueda saber que los cambios en los resultados de interés son provocados por el programa mismo y no por la promoción.
3. La campaña de promoción debe alterar considerablemente las tasas de inscripción en el grupo objeto de la promoción en relación con el grupo que no ha sido objeto de la misma. Normalmente, se piensa en aumentar la inscripción mediante la promoción. Esto se puede verificar constatando que las tasas de inscripción sean más altas en el grupo que es objeto de la promoción que en el grupo que no lo es.

El proceso de promoción aleatoria

Al igual que con los métodos anteriores, se comienza con la población de unidades elegibles para el programa. A diferencia de la asignación aleatoria, ya no se puede elegir aleatoriamente quién recibirá el programa y quién no lo recibirá, porque el programa es totalmente voluntario. Sin embargo, en la población de unidades elegibles, habrá tres tipos de unidades **(Gertler et al., 2017) Ver Figura 6.3:**

- *Siempres.* Aquellos que siempre quieren inscribirse en el programa.
- *Inscritos si se promueve.* Aquellos que se inscriben en el programa solo si son objeto de la promoción.
- *Nuncas.* Aquellos que jamás se inscriben en el programa, independientemente de que sean objeto de la promoción o no.

Figura 6.3. Proceso de promoción aleatoria



Fuente: Gertler et al. (2017) con adaptación propia.

Una vez más, nótese que ser un *Siempre, un Inscrito si se promueve o un Nunca* es una característica intrínseca de las unidades que no se puede medir fácilmente con un equipo de evaluación del programa.

Una vez que se define la *población elegible*, el paso siguiente consiste en seleccionar de manera aleatoria una muestra de la población que formará parte de la evaluación. Estas son las unidades sobre las que se recopilan datos. En algunos casos, por ejemplo, cuando se dispone de datos sobre todas las unidades elegibles, se puede incluir al conjunto de la población en la muestra de evaluación.

Una vez que se ha definido la *muestra de evaluación*, la promoción aleatoria asigna aleatoriamente la muestra mencionada en el grupo objeto de la promoción y un grupo que no es objeto de ella. Dado que tanto los miembros del grupo con promoción como los del grupo sin promoción se escogen de forma aleatoria, ambos grupos compartirán las características de la muestra de evaluación general, que serán equivalentes a las características de la población de las unidades elegibles. Por lo tanto, el grupo que es objeto de la promoción y el grupo que no lo es tendrán características similares.

Después de acabar la campaña de promoción, pueden observarse las tasas de inscripción en ambos grupos. En el grupo sin promoción, se inscribirán solo los *Siempres*. Aunque se sabe qué unidades son *Siempres* en el *grupo sin promoción*, en este grupo no se podrá distinguir entre los *Nuncas y los Inscritos* si se promueve. En cambio, en el grupo con promoción se inscribirán tanto los Inscritos si se promueve como los *Siempres*, mientras que los *Nuncas* no se anotarán. Por ello, en el grupo con promoción se podrá identificar al grupo de *Nuncas*, pero no se podrá distinguir entre los *Inscritos* si se promueve y los *Siempres*.

Estimación de impacto bajo la promoción aleatoria

Suponga que en un grupo de **10** individuos la campaña de promoción aumenta la inscripción de un **30%** en el grupo sin promoción (**3** *Siempres*) a **un 80%** en el grupo con promoción (**3** *Siempres* y **5** *Inscritos* si se promueve). Supóngase que el resultado promedio de todos los individuos del grupo sin promoción (**10** individuos) es **70**, y que el resultado promedio de los individuos en el grupo con promoción (**10** individuos) es **110**. ¿Cuál sería el impacto del programa?. Ver **Figura 6.3.**

Figura 6.3. Estimación del efecto local promedio del tratamiento bajo la promoción aleatoria

	Grupo con promoción	Grupo sin promoción	Impacto
	Porcentaje inscrito= 80% Media *Y* para los asignados a tratamiento = 110	Porcentaje inscrito = 10% Media *Y* para los no asignados a tratamiento = 70	Δ porcentaje de inscritos = 50% ΔY = ITT = 40 LATE = 40%/50% = 80
Nunca se inscribe	Y Y	Y Y	-------
Inscrito si es objeto de la promoción	XXXXXX XXXXXX	XXXXXX XXXXXX	XXXXXX XXXXXX
Siempre se inscribe	Z Z Z	Z Z Z	-------

Nota: Las figuras que aparecen con el fondo sombreado corresponden a los que se inscriben. Δ = impacto causal; Y = resultado.

Fuente: Gertler et al. (2017).

Así:

- En primer lugar, calcúlese la simple diferencia de los resultados entre los grupos con promoción y sin promoción, que es de **40 (110 - 70).** Se sabe que ningún elemento de esta diferencia de **40** proviene de los Nuncas porque estos no se inscriben en ningún grupo. También se sabe que ningún elemento de la diferencia de **40** se debe a los Siempres porque estos se inscriben en ambos grupos. Por lo tanto, toda la diferencia de **40** tendría que deberse al grupo Inscritos si se promueve.
- El segundo paso consiste en obtener la estimación **LATE** del programa de los Inscritos si se promueve. Se sabe que la diferencia de **40** entre los grupos con promoción y sin promoción puede atribuirse a los Inscritos si se promueve, que constituyen solo el **50%** de la población. Para evaluar el efecto promedio del programa en un cumplidor, se divide **40** por el porcentaje de Inscritos si se promueve en la población. Aunque no se puede identificar directamente a los Inscritos si se promueve, se puede deducir cuál debe ser su porcentaje de la población, es decir, la diferencia en las tasas de inscripción de los grupos con promoción y sin promoción **(50%, o 0.5).** Por lo tanto, la estimación del efecto local promedio del tratamiento del programa del grupo Inscritos si se promueve es **40/0.5 = 80.**

Dado que la promoción se asigna de forma aleatoria, los grupos con promoción y sin promoción tienen iguales características. Por lo tanto, las diferencias que se observan en los resultados promedio entre los dos grupos tienen que deberse al hecho de que en el grupo con promoción los Inscritos si se promueve se inscriben, mientras que en el grupo sin promoción no lo hacen. Una vez más, los impactos estimados de los Inscritos si se promueve no deberían extrapolarse directamente a otros grupos, puesto que es probable que sean bastante diferentes de los grupos que se inscriben *Nunca* y *Siempre*.

Limitaciones del método de promoción aleatoria

La promoción aleatoria es una estrategia útil para evaluar el impacto de programas voluntarios y programas con elegibilidad universal, sobre todo porque no requiere la exclusión de ninguna unidad elegible. Sin embargo, el enfoque tiene algunas limitaciones en comparación con la asignación aleatoria del tratamiento **(Gertler et al., 2017):**

- En primer lugar, la estrategia de promoción debe ser efectiva. Si la campaña de promoción no aumenta la inscripción, no aparecerá ninguna diferencia entre los grupos con promoción y sin promoción, y no habrá nada que comparar. Por lo tanto, es crucial diseñar cuidadosamente la campaña de promoción y realizar una prueba piloto extensiva de la misma para asegurarse de que será efectiva. El aspecto positivo es que el diseño de dicha campaña puede ayudar a los administradores del programa enseñándoles cómo aumentar la inscripción después de que haya concluido el período de evaluación.
- En segundo lugar, el método de promoción aleatoria permite estimar el impacto del programa solo para un subconjunto de la población de unidades elegibles (un **LATE**). Concretamente, el impacto promedio local del programa se estima a partir del grupo de individuos que se inscriben únicamente cuando se les incentiva a hacerlo. Sin embargo, puede que los individuos de este grupo tengan características muy diferentes de aquellos que siempre se inscriben o nunca se inscriben. Por lo tanto, el efecto promedio del tratamiento para el conjunto de la población puede ser distinto del efecto del tratamiento promedio estimado para los individuos que participan solo cuando se les incentiva. Una evaluación con promoción aleatoria no estimará los impactos en el grupo de individuos que se inscriben en el programa sin ser incentivados. En algunos casos, este grupo (los *Siempres*) puede ser precisamente el grupo que el programa está diseñado para beneficiar. En este contexto, el diseño de promoción aleatoria arrojará luz sobre los impactos esperados en nuevas poblaciones que se inscribirían debido a la promoción adicional, pero no en cuanto a la población que ya se ha inscrito por su propia iniciativa.

Lista de verificación: promoción aleatoria como variable instrumental

La promoción aleatoria genera estimaciones válidas del contrafactual si la campaña de promoción aumenta de forma considerable la aceptación del programa sin influir directamente en los resultados de interés **(Gertler et al., 2017):**

- Las características de línea de base, ¿están equilibradas entre las unidades que recibieron la campaña de promoción y aquellas que no la recibieron? Compárense las características de línea de base de los dos grupos.
- La campaña de promoción, ¿ha influido de forma considerable en la aceptación del programa? Tendría que influir. Compárense las tasas de aceptación del programa en las submuestras con promoción y sin promoción.
- La campaña de promoción, ¿influye directamente en los resultados? No tendría que influir. Esto no puede comprobarse directamente, de modo que tiene que depender de la teoría, del sentido común y del conocimiento adecuado del entorno de la evaluación de impacto como guía.

Ejemplo utilizando STATA

Otra forma de medir el impacto del programa, cuando el tratamiento no se asigna aleatoriamente, es mediante el método de la *variable instrumental* (**IV**.*Instrumental Variable*). La estimación **IV,** considera la variable de tratamiento (en este caso, participación en programas **pei** de acceso a microcréditos) como *endógeno*. La idea es encontrar una variable o variables *exógenas* observables (instrumentos) que influyan en la variable de participación, pero no influyen en el resultado del programa si participa. Por lo tanto, se querría al menos un instrumento que no esté en las covaria-

bles y eso satisface los requisitos anteriores. La estimación del método **IV** es de doble proceso:

- Primero, la variable de tratamiento se ejecuta contra todas las *covariables*, incluidos los instrumentos.
- Como segunda etapa, se utiliza el valor predicho del tratamiento, en lugar del valor real.

Uso del comando ivreg

El primer paso en la implementación **IV** es *encontrar un instrumento.* En el ejemplo, un la opción de una **pbt** de participar en el programa de microcrédito se utiliza como la **IV**. La elección del **pbt** depende de dos factores: disponibilidad del **pei** de acceso al microcrédito en la región o clúster y la elegibilidad de la **pbt** para participar (que se determina por los activos, instalaciones y recursos técnicos, humanos, finacieros que dedica a las innovaciones. Aunque la colocación del programa en la región o cluster puede ser *endógeno,* la elegibilidad de un **pbt** no lo es, y la combinación de estos dos factores es por lo tanto *exógeno*.

Usando los datos del archivo **pei_08.dta**, cree una variable de programa de para mujeres de la región y, posteriormente, una variable de elección del programa por CEOs de género femenino a nivel de **pbt**. Como se mencionó en ejercicios previos, un **pbt** es elegible para participar en programas de microcrédito si tiene menos de 50 decimales de facilidades o recursos dedicados a la innovación. Por lo tanto:

- Por lo que se deberá teclear (**ver Tabla 6.3**) :

```
use "/Users/DCA/Desktop/STATA log files/pei_08.dta"
egen regionfmf=max(dmmfd), by(region)
gen fchoice=regionfmf==1 & smefac<50
```

Tabla 6.3. Comandos iniciales preparando datos al método IV



```
. use "/Users/DCA/Desktop/STATA log files/pei_01.dta"
(Mexico´s small and medium enterprises based on technology)

. use "/Users/DCA/Desktop/STATA log files/pei_08.dta"

. use "/Users/DCA/Desktop/STATA log files/pei_08.dta"

. egen regionfmf=max(dmmfd), by(region)

. gen fchoice=regionfmf==1 & smefac<50

help unicode_advice

Command
```

Fuente: STATA con datos propios.

- A continuación, *cree instrumentos adicionales* interactuando la variable de selección con todas las covariables. El comando STATA *for* se utiliza para hacerlo en un comando, como se observa (**Ver Tabla 6.4**):

  for var ageceo-educeo lnsmefac bdbaccess pcirr raw1-raw6: gen fchX=fchoice*X

Tabla 6.4. Comando para creación de instrumentos adicionales del método IV



```
. for var ageceo-educeo lnsmefac bdbaccess pcirr raw1-raw6: gen fchX=fchoice*X

-> gen fchageceo=fchoice*ageceo

-> gen fchgenceo=fchoice*genceo

-> gen fcheduceo=fchoice*educeo

-> gen fchlnsmefac=fchoice*lnsmefac

-> gen fchbdbaccess=fchoice*bdbaccess

-> gen fchpcirr=fchoice*pcirr

-> gen fchraw1=fchoice*raw1

-> gen fchraw2=fchoice*raw2

-> gen fchraw3=fchoice*raw3

-> gen fchraw4=fchoice*raw4

-> gen fchraw5=fchoice*raw5

-> gen fchraw6=fchoice*raw6

Command
```

Fuente: STATA con datos propios.

- El paso siguiente, es la implementación **IV**, el cual usa el comando STATA *ivreg*. La primera etapa de la ecuación aparece dentro de paréntesis en la sintaxis y la primera opción se muestra, en los resultados de la primera etapa (**ver Tabla 6.5**):

  ivreg lexptot ageceo-educeo ln smefac bdbaccess pcirr raw1-raw6 (dfmfd = ageceo-educeo lnsmefac bdbaccess pcirr raw1-raw6 fch*), first

Tabla 6.5. Comando y tabulación del método IV en primera etapa

```
. ivreg lexptot ageceo-educeo ln smefac bdbaccess pcirr raw1-raw6 (dfmfd = ageceo-educeo lnsmefac bdbaccess pcirr raw1-raw6 fch*), first
note: fchraw4 dropped because of collinearity

First-stage regressions
```

Source	SS	df	MS			
				Number of obs	=	1,129
				F(24, 1104)	=	5.22
Model	28.6964728	24	1.19568637	Prob > F	=	0.0000
Residual	252.729568	1,104	.22892171	R-squared	=	0.1020
				Adj R-squared	=	0.0824
Total	281.426041	1,128	.249491171	Root MSE	=	.47846

dfmfd	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ageceo	-.0025939	.0019654	-1.32	0.187	-.0064502	.0012625
genceo	-.0783472	.0977156	-0.80	0.423	-.2700766	.1133821
educeo	-.0141505	.0068058	-2.08	0.038	-.0275043	-.0007967
lnsmefac	-.304718	.0946578	-3.22	0.001	-.4904474	-.1189886
smefac	.0001462	.000149	0.98	0.327	-.0001462	.0004386
bdbaccess	-.1041506	.0722404	-1.44	0.150	-.2458946	.0375935
pcirr	-.0572703	.0888351	-0.64	0.519	-.231575	.1170344
raw1	.0042525	.0161359	0.26	0.792	-.0274079	.0359129
raw2	.0046802	.032365	0.14	0.885	-.0588236	.0681841
raw3	-.0292019	.010904	-2.68	0.008	-.0505967	-.0078071
raw4	0	(omitted)				
raw5	.0152968	.095236	0.16	0.872	-.1715673	.2021608
raw6	.001367	.0066186	0.21	0.836	-.0116194	.0143534
fchoice	-1.450499	.5097677	-2.85	0.005	-2.450722	-.450276
fchageceo	.006132	.0024517	2.50	0.013	.0013216	.0109424
fchgenceo	.111294	.1139648	0.98	0.329	-.112318	.334906
fcheduceo	-.0035804	.0090273	-0.40	0.692	-.0212929	.0141322
fchlnsmefac	.237749	.1870316	1.27	0.204	-.1292285	.6047266
fchbdbaccess	.0807825	.0869065	0.93	0.353	-.0897381	.2513032
fchpcirr	.20461	.1064609	1.92	0.055	-.0042785	.4134985
fchraw1	-.0183588	.0200603	-0.92	0.360	-.0577194	.0210018
fchraw2	-.0063326	.0385833	-0.16	0.870	-.0820374	.0693722
fchraw3	.0407077	.0129689	3.14	0.002	.0152612	.0661542
fchraw5	.0880632	.1126137	0.78	0.434	-.1328978	.3090242
fchraw6	.010188	.0080019	1.27	0.203	-.0055126	.0258886
_cons	1.192447	.4290294	2.78	0.006	.3506423	2.034252

Command

Fuente: STATA con datos propios.

- La salida muestra los resultados de la primera etapa primero y luego los resultados de la segunda etapa. **Ver Tabla 6.6.**

Tabla 6.6. Comando y tabulación del método IV en segunda etapa

Instrumental variables (2SLS) regression

Source	SS	df	MS			
Model	62.7964026	13	4.83049251	Number of obs	=	1,129
Residual	235.063527	1,115	.210819307	F(13, 1115)	=	22.16
				Prob > F	=	0.0000
				R-squared	=	0.2108
				Adj R-squared	=	0.2016
Total	297.85993	1,128	.264060221	Root MSE	=	.45915

lexptot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
dfmfd	.0913451	.1650758	0.55	0.580	-.2325491	.4152394
ageceo	.003438	.0011428	3.01	0.003	.0011956	.0056803
genceo	-.0577462	.0479921	-1.20	0.229	-.1519111	.0364188
educeo	.0500758	.0049206	10.18	0.000	.0404211	.0597304
lnsmefac	.1341749	.0577778	2.32	0.020	.0208095	.2475403
smefac	.0001245	.0001152	1.08	0.280	-.0001014	.0003505
bdbaccess	-.0139594	.0397206	-0.35	0.725	-.0918948	.0639761
pcirr	.1566435	.0490024	3.20	0.001	.0604962	.2527908
raw1	.0040245	.0092018	0.44	0.662	-.0140302	.0220793
raw2	-.0390483	.016843	-2.32	0.021	-.0720958	-.0060007
raw3	.0204149	.0056355	3.62	0.000	.0093576	.0314722
raw4	0	(omitted)				
raw5	.114694	.0497246	2.31	0.021	.0171297	.2122582
raw6	.0098954	.0037954	2.61	0.009	.0024484	.0173424
_cons	7.45806	.2220666	33.58	0.000	7.022345	7.893776

Instrumented: dfmfd
Instruments: ageceo genceo educeo lnsmefac smefac bdbaccess pcirr raw1 raw2 raw3 raw4 raw5 raw6 fchoice fchageceo fchgenceo fcheduceo fchlnsmefac fchbdbaccess fchpcirr fchraw1 fchraw2 fchraw3 fchraw5 fchraw6

Command

Fuente: STATA con datos propios

- Según el producto de la primera etapa, la educación del CEO y las los activos de las **pbt** en los activos, recursos técnicos, humanos, financieros dedicados a la innovación influyen negativamente en la participación en el programa de microcrédito; también lo hacen

los instrumentos. Los resultados de la segunda etapa muestran que después de controlar la endogeneidad del programa participación, la participación de las CEOs femeninas de **pbt** en **pei** con acceso a microcrédito tiene un impacto significativo (**9.1 %**) en el gasto per cápita del hogar (***t* = 0.55**).

Prueba de endogeneidad: OLS vs IV

Se pueden usar algunas pruebas para determinar, si un mínimo cuadrado ordinario (**OLS**) o el método por **IV** es más apropiado. STATA tiene un comando *ivendog* que realiza una *prueba F y chi-cuadrado,* siguiendo las metodologías llamadas pruebas de: ***Wu-Hausman*** y la de ***Durbin-Wu-Hausman***, respectivamente. La hipótesis nula, es que OLS es consistente (en este caso, implica, que el tratamiento es *exógeno*) Si no se rechaza la hipótesis nula, una OLS debería ser suficiente; de lo contrario, debe usarse un método **IV**. El comando *ivendog* se usa después de Comando ivreg:

ivendog

Los resultados muestran, que la hipótesis nula se rechaza al nivel del 10 por ciento, *lo que implica que IV es un mejor modelo que OLS* (**Ver Tabla 6.7**)*:*

Tabla 6.7. Comando Ivendog en método IV

```
. ivendog

Tests of endogeneity of: dfmfd
H0: Regressor is exogenous
    Wu-Hausman F test:                0.00002   F(1,1114)    P-value = 0.99611
    Durbin-Wu-Hausman chi-sq test:    0.00002   Chi-sq(1)    P-value = 0.99608

.
Command
```

Fuente: STATA con datos propios.

Método IV por tratamiento binario: comando treatreg

Los métodos de estimación **IV** anteriores, se aplican cuando el *regresor endógeno es continuo*. Cuando el *regresor endógeno es binario* (participante / no participante), utilizando un modelo lineal en la primera etapa del procedimiento **IV** puede o no ser apropiado. Otro método que se ajusta a un modelo de efectos-tratamiento cuando el regresor endógeno es binario es el comando **STATA** *treatreg*. El comando *treatreg* se ajusta amodelo efectos-tratamiento utilizando, la *máxima probabilidad completa* (*full maximum likelihood*) o el estimador consistente de dos pasos (*two-step consistent estimator*).

El comando *treatreg* tiene en cuenta, el efecto de la variable endógena binaria (*binary endogenous variable)* sobre el resultado de intereses condicionado a los dos conjuntos de variables exógenas. El comando estima dos regresiones simultáneamente:

- La primera ecuación es estimada mediante *regresión probit* para predecir la probabilidad de tratamiento.
- El segundo es ya sea una *regresión lineal o probit* para las variables de resultado.

Los dos términos de error, se supone que se distribuyen normalmente, de manera conjunta.

Para aplicarlo:

- Teclear comando *teatreg* (**ver Tabla 6.8**):

  treatreg lexptot ageceo-educeo smefac bdbaccess pcirr raw1-raw2, treat (dfmfd= ageceo-educeo smefac bdbaccess pcirr raw1-raw2 fch*)

Tabla 6.8. Comando Treateg

```
. treatreg lexptot ageceo-educeo smefac bdbaccess pcirr raw1-raw2, treat (dfmfd= ageceo-educeo smefa
> c bdbaccess pcirr raw1-raw2 fch*)
note: fchraw4 omitted because of collinearity

Iteration 0:   log likelihood = -1463.8163
Iteration 1:   log likelihood = -1461.3713
Iteration 2:   log likelihood = -1460.4267
Iteration 3:   log likelihood = -1460.4199
Iteration 4:   log likelihood = -1460.4199

Treatment-effects model -- MLE                  Number of obs     =      1,129
                                                Wald chi2(9)      =     224.42
Log likelihood = -1460.4199                     Prob > chi2       =     0.0000
```

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lexptot						
ageceo	.0039372	.0012459	3.16	0.002	.0014954	.0063791
genceo	-.0528138	.0543079	-0.97	0.331	-.1592553	.0536276
educeo	.060341	.0049787	12.12	0.000	.0505829	.0700991
smefac	.0005595	.0000857	6.53	0.000	.0003915	.0007276
bdbaccess	.0253525	.0435293	0.58	0.560	-.0599635	.1106684
pcirr	-.0065774	.0483582	-0.14	0.892	-.1013577	.0882029
raw1	.0163472	.0100791	1.62	0.105	-.0034075	.0361018
raw2	-.0431614	.0186765	-2.31	0.021	-.0797666	-.0065561
dfmfd	.5647996	.0901804	6.26	0.000	.3880493	.7415498
_cons	7.973683	.1889165	42.21	0.000	7.603413	8.343952
dfmfd						
ageceo	-.0071907	.0050549	-1.42	0.155	-.0170981	.0027167
genceo	-.1643643	.2559274	-0.64	0.521	-.6659924	.3372638
educeo	-.0562197	.0182216	-3.09	0.002	-.0919334	-.020506
smefac	-.0015681	.0004912	-3.19	0.001	-.0025308	-.0006054
bdbaccess	-.4638754	.1783876	-2.60	0.009	-.8135006	-.1142422
pcirr	.1959245	.1969537	0.99	0.320	-.1900977	.5819467
raw1	-.0195237	.0395763	-0.49	0.622	-.0970917	.0580444
raw2	.0209449	.080686	0.26	0.795	-.1371967	.1790865
fchoice	-3.881027	1.016217	-3.82	0.000	-5.872776	-1.889279
fchageceo	.0139302	.0060919	2.29	0.022	.0019903	.0258701
fchgenceo	.2103183	.2862926	0.73	0.463	-.3508048	.7714415
fcheduceo	.0197889	.0225672	0.88	0.381	-.0244421	.0640199
fchlnsmefac	.0173663	.4038339	0.04	0.966	-.7741335	.8088661
fchbdbaccess	.4271934	.2067795	2.07	0.039	.0219131	.8324738
fchpcirr	.3057412	.2360096	1.30	0.195	-.156986	.7684684
fchraw1	-.0182424	.0472589	-0.39	0.699	-.1108524	.0743676
fchraw2	-.0081228	.0928215	-0.09	0.930	-.1900496	.1738041
fchraw3	.0487262	.0168293	2.90	0.004	.0157413	.0817111
fchraw4	0	(omitted)				
fchraw5	.3584211	.1436007	2.50	0.013	.0769689	.6398733
fchraw6	.0384589	.0104409	3.68	0.000	.017995	.0589227
_cons	1.103378	.8103061	1.36	0.173	-.484793	2.691549
/athrho	-.6667375	.1260692	-5.29	0.000	-.9138286	-.4196463
/lnsigma	-.6541054	.0427001	-15.32	0.000	-.7378762	-.5704947
rho	-.5828297	.0832447			-.7229648	-.3966324
sigma	.5198654	.0221983			.4781283	.5652457
lambda	-.302993	.0549082			-.4105954	-.1953985

```
Wald test of indep. eqns. (rho = 0): chi2(1) =    27.97   Prob > chi2 = 0.0000
```

Command

Fuente: STATA con datos propios.

- El resultado mostrado en la **Tabla 6.8** es el método de efecto del tratamiento utilizando la estimación de *máxima verosimilitud*. Muestra que la participación de las CEO femeninas tiene un impacto significativo positivo en la **pbt**. **(t = 6.26).**

Método IV con efectos fijos: estudios transversales

La regresión de variables instrumentales (**IV**) se puede combinar con *efectos fijos*. Aquí una demostración utilizando datos de transversales de muestra. El comando a usar es *xtivreg* con la opción *fe*, por lo que:

- Se ejecuta una *regresión de efectos fijos* a nivel de región utilizando el mismo **pei_08.dta**. Aquí está el comando para la participación de las CEO femeninas de **pbt** que participan en el **pei** de acceso a microcrédito:

  xtivreg lexptot year ageceo-educeo lnsmefac bdbaccess pcirr raw1-raw6 (dfmfd= ageceo-educeo lnsmefac bdbaccess pcirr raw1-raw6 fch*), fe i(region)

- A continuación, ejecute una regresión de *efectos fijos* a nivel de región con el mismo **pei_08.dta**. *Utilizando los efectos fijos a nivel de región se hacen que desaparezcan, los impactos de participación* (**ver Tabla 6.9**):

Tabla 6.9. Comandos y tabulación método IV con efectos fijos para estudios transversales

```
. xtivreg lexptot year ageceo-educeo lnsmefac bdbaccess pcirr raw1-raw6 (dfmfd= ageceo-educeo lnsmef
> ac bdbaccess pcirr raw1-raw6 fch*), fe i(region)
note: year omitted because of collinearity
note: raw4 omitted because of collinearity
note: raw4 omitted because of collinearity
note: fchraw4 omitted because of collinearity

Fixed-effects (within) IV regression          Number of obs      =      1,129
Group variable: region                        Number of groups   =          4

R-sq:                                         Obs per group:
     within  = 0.2104                                        min =         89
     between = 0.3600                                        avg =      282.2
     overall = 0.2095                                        max =        338

                                              Wald chi2(12)      =  384094.96
corr(u_i, Xb)  = -0.0039                      Prob > chi2        =     0.0000
```

lexptot	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
dfmfd	.0830376	.1655206	0.50	0.616	-.2413768	.4074519
year	0	(omitted)				
ageceo	.0031131	.0011433	2.72	0.006	.0008723	.0053539
genceo	-.0593601	.0482394	-1.23	0.218	-.1539075	.0351874
educeo	.0498416	.0048587	10.26	0.000	.0403187	.0593646
lnsmefac	.1789871	.0435476	4.11	0.000	.0936354	.2643388
bdbaccess	-.005659	.0396703	-0.14	0.887	-.0834113	.0720933
pcirr	.1479055	.0487655	3.03	0.002	.0523269	.2434842
raw1	-.0012131	.0093961	-0.13	0.897	-.0196291	.017203
raw2	-.0413303	.017343	-2.38	0.017	-.075322	-.0073385
raw3	.0196279	.0056699	3.46	0.001	.008515	.0307407
raw4	0	(omitted)				
raw5	.123496	.0493441	2.50	0.012	.0267833	.2202087
raw6	.0101162	.0037396	2.71	0.007	.0027868	.0174455
_cons	7.523545	.2243061	33.54	0.000	7.083913	7.963177
sigma_u	.06239284					
sigma_e	.45833259					
rho	.01819423	(fraction of variance due to u_i)				

```
F  test that all u_i=0:      F(3,1113) =      2.40           Prob > F     = 0.0660

Instrumented:   dfmfd
Instruments:    ageceo genceo educeo lnsmefac bdbaccess pcirr raw1 raw2 raw3
                raw5 raw6 fchoice fchageceo fchgenceo fcheduceo fchlnsmefac
                fchbdbaccess fchpcirr fchraw1 fchraw2 fchraw3 fchraw5 fchraw6
```

Fuente: STATA con datos propios.

Método IV con efectos fijos: estimados por panel

Ahora se muestra una implementación de *xtivreg* utilizando los datos del panel de **pei_0108.dta**. Después de crear las variables necesarias como antes, emita el comando *xtivreg* (**Ver Tabla 6.10**)

```
use "/Users/DCA/Desktop/STATA log files/pei_0108.dta"
gen lnsmefac=ln(1+smefac/100)
gen lexptot=ln(1+exptot)
egen regionfmf=max(dmmfd), by(region)
gen fchoice=regionfmf==1 & smefac<50
for var ageceo-educeo lnsmefac bdbaccess pcirr raw1-raw6: gen fchX
fchoice*X xtivreg lexptot year ageceo-educeo lnsmefac bdbaccess
pcirr raw1-raw6 (dfmfd= ageceo-educeo lnsmefac bdbaccess pcirr
raw1-raw2 fch*), fe i(id)
```

Tabla 6.10. Comandos con efectos fijos: estimados panel



Fuente: STATA con datos propios.

Los resultados no muestran alguna participación de impacto en el gasto de las **pbt. Ver Tabla 6.11**

Tabla 6.11. Comandos y tabla de resultados con efectos fijos: estimados panel

```
. xtivreg lexptot year ageceo-educeo lnsmefac bdbaccess pcirr raw1-raw6 (dfmfd= ageceo-educeo lnsmefac bdba
> ccess pcirr raw1-raw2 fch*), fe i(id)

Fixed-effects (within) IV regression            Number of obs     =      1,652
Group variable: id                              Number of groups  =        826

R-sq:                                           Obs per group:
     within  = 0.1639                                         min =          2
     between = 0.1772                                         avg =        2.0
     overall = 0.1642                                         max =          2

                                                Wald chi2(14)     =  863895.89
corr(u_i, Xb)  = 0.1058                         Prob > chi2       =     0.0000

------------------------------------------------------------------------------
     lexptot |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       dfmfd |   .0890017   .2170716     0.41   0.682    -.3364508    .5144542
        year |    .225843   .0811626     2.78   0.005     .0667673    .3849188
      ageceo |  -.0001554   .0018938    -0.08   0.935    -.0038671    .0035564
      genceo |  -.0557328   .0732937    -0.76   0.447    -.1993857    .0879202
      educeo |   .0141314   .0086335     1.64   0.102      -.00279    .0310528
    lnsmefac |   .1326746   .0645482     2.06   0.040     .0061625    .2591867
   bdbaccess |  -.1007024   .0541365    -1.86   0.063    -.2068081    .0054032
       pcirr |   .0942596   .0645709     1.46   0.144    -.0322971    .2208163
        raw1 |   .0206327   .0134302     1.54   0.124    -.0056899    .0469554
        raw2 |   -.025712   .0136677    -1.88   0.060    -.0525002    .0010762
        raw3 |  -.0027496   .0065599    -0.42   0.675    -.0156069    .0101076
        raw4 |  -.0205451   .0167487    -1.23   0.220    -.0533719    .0122817
        raw5 |   .1721962    .048939     3.52   0.000     .0762774    .2681149
        raw6 |   .0045124   .0031632     1.43   0.154    -.0016872    .0107121
       _cons |    7.83581   .2521263    31.08   0.000     7.341652    8.329968
-------------+----------------------------------------------------------------
     sigma_u |  .34741977
     sigma_e |  .36531242
         rho |   .4749115   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```


```
     sigma_u |  .34741977
     sigma_e |  .36531242
         rho |   .4749115   (fraction of variance due to u_i)
```


```
------------------------------------------------------------------------------
 F  test that all u_i=0:     F(825,812) =     1.57         Prob > F    = 0.0000
------------------------------------------------------------------------------
Instrumented:   dfmfd
Instruments:    year ageceo genceo educeo lnsmefac bdbaccess pcirr raw1 raw2
                raw3 raw4 raw5 raw6 fchoice fchageceo fchgenceo fcheduceo
                fchlnsmefac fchbdbaccess fchpcirr fchraw1 fchraw2 fchraw3
                fchraw4 fchraw5 fchraw6
------------------------------------------------------------------------------

.

Command
```

Fuente: STATA con datos propios.

CAPÍTULO 7.
Regresión Discontínua (*RD. Regression Discontinuity*)

Los programas sociales a menudo utilizan un índice para decidir quién tiene derecho a inscribirse en un programa y quién no. Por ejemplo, los programas de lucha contra la pobreza o incentivos a la innovación suelen focalizarse en los en pymes u hogares pobres, identificados mediante una puntuación o un índice de baja competitividad o de pobreza. Dichos índices se suelen en una fórmula que mide un conjunto de activos básicos de la pyme como factor aproximado (o estimativo) de sus medios (como el ingreso, el consumo o el poder adquisitivo). Se clasifica a las pymes con baja puntuación como de baja competitividad, y a las pymes con puntuaciones más altas se les considera relativamente acomodados. Los programas de lucha contra el bajo índice de innovación o competencia, suelen establecer un umbral o una puntuación límite, por debajo del cual se determina la condición de baja capacidad de innovación o pobreza y la elegibilidad para el programa. **(Gertler et al., 2017)**

Discontinuidades y demoras en la implementación del programa, según los *criterios de elegibilidad* o otros *factores exógenos*, pueden ser muy útiles en la evaluación de *programas no experimentales*. Los sujetos, por encima y por debajo del umbral, suponiendo que sean similares en las características observadas, se pueden distinguir en términos de resultados. Sin embargo, las muestras, a través de las cuales se compara, tendrían que estar lo suficientemente cerca del *límite de elegibilidad* para garantizar la comparabilidad. Además, la *heterogeneidad no observada*, puede ser un factor si los sujetos dentro del rango de *selección elegible* muestran una variación en la implementación real del programa, lo que lleva a *sesgo de selección*. En ese caso, muestras *elegibles* y *no elegibles* cercanas al *límite de elegibilidad* se tomarían para comparar el *efecto promedio del*

programa. Los enfoques de discontinuidad son, por lo tanto, similares a los métodos de variables instrumentales (**IV**) porque introducen una variable exógena que está altamente correlacionada con la participación, aunque no es similar a la participación **(Khandker et al., 2017).**

El diseño de regresión discontinua (**RD.** *Regression Discontinuity*) es un método de evaluación de impacto que se puede utilizar en programas que tienen un índice de elegibilidad continuo con un umbral (puntuación límite) de elegibilidad definido con claridad para determinar quién es elegible y quién no lo es. A fin de aplicar un **RD**, deben cumplirse las siguientes condiciones **(Gertler et al., 2017):**

1. El índice debe clasificar a las personas o unidades de una manera continua o fluida. Índices como el de baja innvación o pobreza, las puntuaciones de las pruebas estandarizadas o los ingresos de la pyme, la edad tienen numerosos valores que se pueden ordenar de menor a mayor y, por lo tanto, se pueden considerar continuos. En cambio, las variables con categorías discretas que solo tienen unos pocos valores posibles o no se pueden ordenar, no se consideran continuas. Ejemplos de esta última clase son la condición laboral (empleado o desempleado), el nivel más alto de estudios alcanzado (primario, secundario, universitario o posgrado), la propiedad de un automóvil (sí o no) o el país de nacimiento.

2. El índice debe tener una puntuación límite claramente definida, es decir, un punto por debajo o por encima del cual se clasifica a la población como elegible para el programa. Por ejemplo, las **pbt** con un índice de innovación igual o menor a **50** sobre **100** se podrían clasificar como de bajo índice, las **pbt** con cierto nivel de ingresos anuales por innovaciones con cierta cantidad de empleados, se podrían clasificar como elegibles para los incentivos **pei**. Esto establecería puntuaciones límite en cada uno de los rubros, respectivamente.

3. La puntuación límite debe ser única para el programa de interés, es decir, aparte del programa que se evalúa, no debería haber otros

programas que utilicen la misma puntuación límite. Por ejemplo, si un índice de de baja índice de innovación por debajo de 50 clasifica a una **pbt** para recibir un incentivo, no se podría utilizar el método **RD** para estimar por sí solo el impacto del programa de transferencias de efectivo.

4. La puntuación de un sujeto o una unidad particular no puede ser manipulada por los encuestadores, los beneficiarios potenciales, los administradores del programa o los políticos.

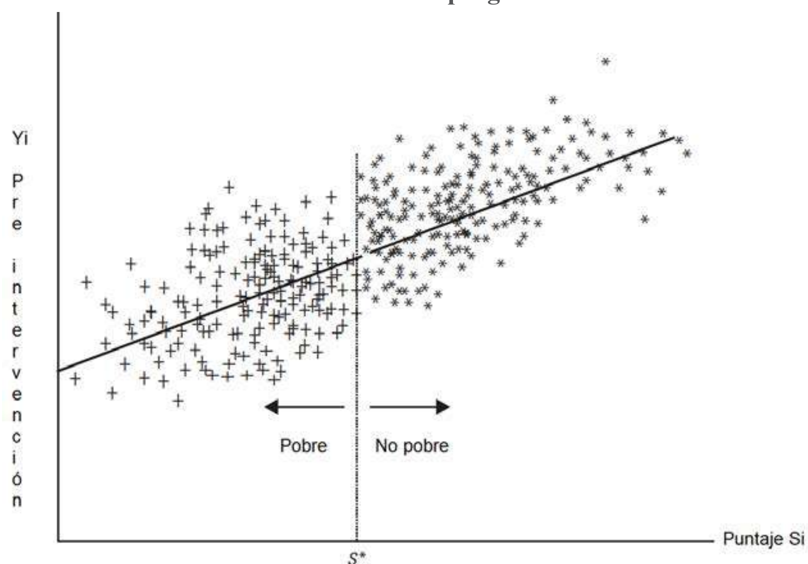
El **RD** estima el impacto en torno a la puntuación límite de elegibilidad como la diferencia entre el resultado promedio de unidades del lado tratado de la puntuación límite de elegibilidad y el resultado promedio de unidades en el lado no tratado (comparación) de la puntuación límite.

Teoría de la regresión de discontinuidad

Como se ha venido discutiendo, algunos diseños no experimentales de asignación de política de innovación proveen fuentes de variación del tratamiento que pueden ser explotadas para estimar el efecto del tratamiento bajo supuestos relativamente débiles. Un caso especial de este conjunto de diseños ocurre cuando la probabilidad de participación en el tratamiento cambia díscantinuamente con una variable continua observada ***Z.*** Esto ocurre, generalmente, en casos en los cuales la política de innovación se focaliza en un grupo particular de la sujetos con base en un instrumento de focalización que es una variable continua ***Z*** y que el investigador puede observar. La literatura de **RD** se ha centrado al momento en el caso del instrumento de focalización ***Z*** contínuo. Los trabajos de los últimos años, presentan diseños **RD** con variable **Z** *discreta* y *difusa*. En este caso, la asignación al tratamiento depende completa o parcialmente del valor de la variable de focalización, que llamamos **Z**. En particular, la asignación al tratamiento se determina con base en un *umbral de* ***Z***, que determina la elegibilidad para el programa **(Peña y Bernal, 2011).**

Para modelar el efecto de un programa particular, en los resultados individuales y_i a través del enfoque **RD**, se requiere una variable S_i que determine la elegibilidad del programa (tales como el tamaño de la empresa, el género del CEO , la edad del CEO, etc.) con un *corte de elegibilidad* de ***s*.*** La ecuación de estimación es $y_i = \beta S_i + \varepsilon_i$ donde los sujetos con $s_i <= s^*$ por ejemplo, reciben el programa, y los sujetos s_i con $> s^*$ no son elegibles para participar. Los sujetos, en una banda estrecha arriba y debajo de ***s**** necesitan ser *comparables* en el sentido de que se espera, que logren resultados similares antes del programa de intervención. La **Gráfica 7.1** muestra un ejemplo de esta propiedad, donde los sujetos por debajo de ***s*,*** se consideran pobres, y aquellos por encima del umbral, se consideran no pobres.

Gráfica 7.1. Resultados antes del programa de intervención



Fuente: Khandker et al. (2017) con adaptación propia.

Si se supone, que existen límites a ambos lados del umbral ***s****, el estimador de impacto para un **ε> 0** arbitrariamente pequeño, alrededor del umbral, sería el siguiente **(Khandker et al., 2017)**:

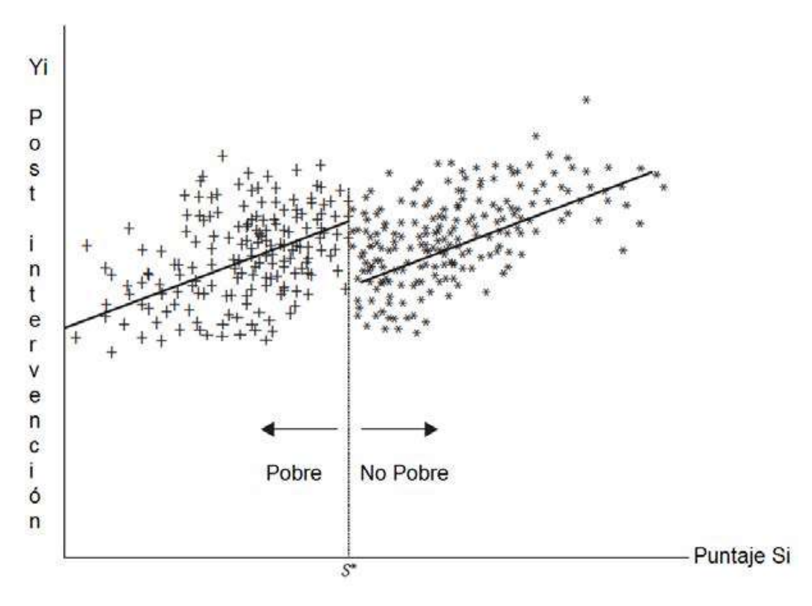
$$E [y_i \mid s^* - \varepsilon] - E [y_i \mid s^* + \varepsilon] = E [\beta S_i \mid s^* - \varepsilon] - E [\beta S_i \mid s^* + \varepsilon] \ldots (7.1)$$

Llevando al límite ambos lados de la **ecuación 7.1** como $\varepsilon \rightarrow 0$ identificando $\boldsymbol{\beta}$ como la relación de la diferencia en los resultados de los individuos, justo por encima y por debajo del umbral, ponderado por la diferencia en sus realizaciones de $\boldsymbol{S_i}$:

$$\lim_{\varepsilon \to 0} E[y_i \mid s^* - \varepsilon] - \lim_{\varepsilon \to 0} E[y_i \mid s^* + \varepsilon] = y^- - y^+ = \beta (S^- - S^+)$$
$$\Rightarrow \beta = \frac{y^- - y^+}{S^- - S^+} \ldots\ldots\ldots\ldots\ldots (7.2)$$

De acuerdo con la configuración en la **Gráfica 7.1**, los resultados después de la intervención del programa, medidos por el modelo de discontinuidad, se reflejan en la **Gráfica 7.2.**

Gráfica 7.2. Resultados después del programa de intervención



Fuente: Khandker et al. (2017) con adaptación propia.

Debido a que en la práctica, la determinación o ejecución de la elegibilidad, puede no ser *agudo* (como en un *experimento aleatorio*), ***s*** se puede reemplazar con una probabilidad de participar **P (*S*) = E (*T* | *S*),** donde ***T* = 1** si se recibe tratamiento y ***T* = 0** en caso contrario **(Ravallion, 2008).** En este caso, la discontinuidad es estocástica o *difusa*, y en lugar de medir las diferencias en los resultados anteriores, arriba y abajo de de ***s*,*** el estimador de impacto mediría la diferencia alrededor de un vecindario de ***s****. Este resultado puede ocurrir cuando las *reglas de elegibilidad* no se cumplen estrictamente o cuando ciertas áreas geográficas están focalizadas, pero los límites no están bien definidos y la movilidad es común. Si el *umbral de elegibilidad*, es determinado de manera exógena por el programa y altamente correlacionado con el tratamiento, también se podría usar **s*** como un **IV** para la participación.

El diseño de regresión discontinua difuso

Aun cuando se haya verificado que no existe evidencia de manipulación en el *índice de elegibilidad*, puede que todavía persista un problema si las unidades no respetan su asignación al grupo de tratamiento o de comparación. En otras palabras, algunas unidades que tienen derecho al programa sobre la base de su índice de elegibilidad pueden decidir no participar, mientras que otras unidades que no tenían derecho al programa sobre la base de su índice de elegibilidad pueden encontrar un modo de participar de todas maneras. Cuando todas las unidades cumplen con la asignación que les corresponde sobre la base de su índice de elegibilidad, se dice que el **RD** es *nítido*, y si hay incumplimiento en alguno de los lados de la puntuación límite, así, el **RD** es *difuso* (**Gráfica 7.3**).

Gráfica 7.3. RD difuso vs. nítido

Fuente: Elaboración propia.

Si el **RD** es *difuso*, se puede utilizar el enfoque de variable instrumental para corregir por la falta de cumplimiento (véase el capítulo 6). Recuérdese que en el caso de la asignación aleatoria con incumplimiento, se utiliza la asignación aleatoria como la variable instrumental que ayudó a corregir por la falta de cumplimiento. En el caso del **RD**, se puede usar la asignación original basada en el índice de elegibilidad como variable instrumental. Sin embargo, hacerlo tiene un inconveniente, a saber, que la estimación de impacto con el **RD** instrumental será más localizada en el sentido de que ya no es válida para todas las observaciones cercanas a la puntuación límite sino que representa el impacto para el subgrupo de la población situada cerca de la puntuación límite y que participa en el programa solo debido a los criterios de elegibilidad.

Pasos para aplicar RD

La regresión no paramétrica estándar, se puede utilizar para estimar el efecto del tratamiento, ya sea la configuración *de discontinuidad de regresión aguda* o *difusa*. Para un *diseño agudo de discontinuidad*, el efecto del tratamiento puede estimarse mediante una simple comparación de los resultados medios de los sujetos a la dereha e izquierda del umbral. Espe-

cíficamente, las regresiones lineales locales sobre el resultado ***y***, dado un conjunto de covariables ***x*, deben** ejecutarse para sujetos a ambos lados del umbral, para estimar la diferencia **(Khandker et al., 2017):**

$$y^- - y^+ = \lim_{s_i \uparrow s*} E(y_i | s_i = s*) - \lim_{s_i \downarrow s*} E(y_i | s_i = s*), \quad \text{.................}(7.3)$$

Como un ejemplo, y^- *e* y^+_i puede especificarse a través de estimaciones Kernel:

$$y^- = \frac{\sum_{i=1}^{n} y_i \alpha_i^* K(u_i)}{\sum_{i=1}^{n} \alpha_i^* K(u_i)}$$

$$y^+ = \frac{\sum_{i=1}^{n} y_i (1-\alpha_i)^* K(u_i)}{\sum_{i=1}^{n} (1-\alpha_i)^* K(u_i)} \quad \text{……………}(7.4)$$

Para un *diseño de discontinuidad difusa*, se necesita un proceso de dos pasos:

1. La *regresión lineal*, se puede aplicar en el resultado para sujetos en ambos lados del umbral para determinar la magnitud de la diferencia (*o discontinuidad*) para el resultado.
2. Similarmente a la *regresión lineal* se puede aplicar al indicador de tratamiento para llegar a una diferencia o *discontinuidad* para el indicador de tratamiento. El radio del resultado de discontinuidad, al tratamiento es el efecto del tratamiento para una discontinuidad difusa diseño. El radio del resultado de la discontinuidad, al *tratamiento de la discontinuidad,* es el efecto del tratamiento para un diseño de discontinuidad difusa.

Aunque los impactos, en un vecindario del *punto de corte* no son identificados como paramétricos en diseños de discontinuidad, la literatura apli-

cada ha utilizado con mayor frecuencia, una alternativa de método paramétrico en el que se utiliza la discontinuidad en el criterio de elegibilidad como un **IV** para la colocación del programa.

Graficar el efecto del tratamiento previsto, también proporciona un contraste útil entre las poblaciones elegibles y no elegibles, así como para aquellos qu ese ubican en una banda estrecha alrededor del límite. Trazar la densidad de la variable, que determina la elegibilidad alrededor del el umbral, también puede ayudar a mostrar si el diseño de **RD** es válido, es decir, que los miembros de la muestra no elegible en última instancia no se convierten en participantes, lo que podría suceder, por ejemplo, si conocían el umbral y ajustaban su valor reportado, de la variable de elegibilidad para calificar). Trazar los valores promedio de las covariables, alrededor del umbral, también puede proporcionar una indicación de problemas de especificación, ya sea con una *discontinuidad aguda* o un enfoque de *discontinuidad difusa*.

Variaciones de RD

Son posibles numerosas variaciones de diseños de **RD**, incluidas las combinaciones con experimentos de desempate aleatorizado, alrededor del umbral para crear inferencias más fuertes. Así, se tiene **(Khandker et al., 2017)**:

1. Dependiendo de la naturaleza de la *regla de elegibilidad*, es decir, si es una función de un cambio variable a lo largo del tiempo o si proviene de una intervención única, panel o sección transversal, los datos se pueden usar en el análisis **RD**. Una *aleatorización de desempate*, por ejemplo, involucra una situación donde una superposición ocurre entre los grupos de tratamiento y control en la variable que determina la elegibilidad para el programa. En esta situación, el tratamiento se asignaría aleatoriamente a las observaciones en el área de superposición. La **Gráfica 7.4** describe esta situación.

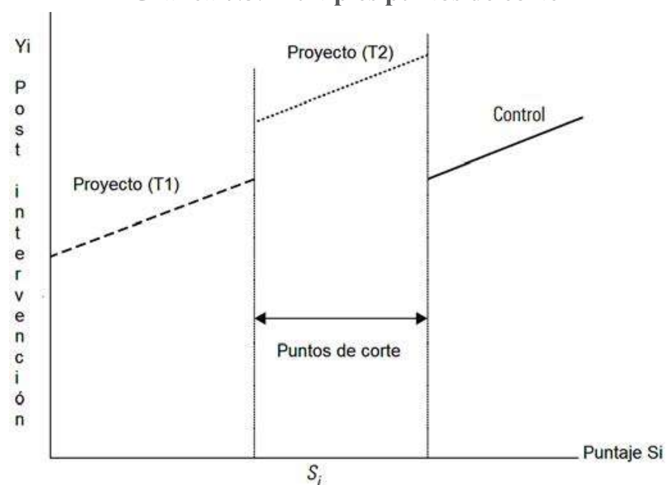
Gráfica 7.4. Usando un experimento de desempate

Asignación aleatoria
Yi
Post intervención
Proyecto
Control
Puntaje Si
S_i

Fuente: Khandker et al. (2017) con adaptación propia.

2. Otra variante, es donde se puede explotar más de un punto de corte para comparar los efectos del tratamiento La regresión correspondiente para estimar el impacto del programa, por lo tanto, incluiría dos grupos de tratamiento, uno correspondiente a cada discontinuidad. La Gráfica 7.5 describe este contexto.

Gráfica 7.5. Multiples puntos de corte

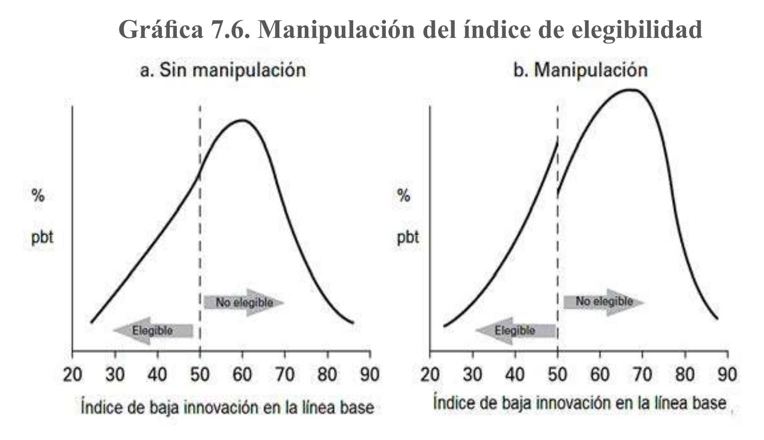


Fuente: Khandker et al. (2017) con adaptación propia.

Verificación de la validez del diseño de RD

Para que un **RD** produzca una estimación **LATE** no sesgada de la puntuación límite, es importante que el índice de elegibilidad no sea manipulado en la cercanía de la puntuación límite de modo que un individuo pueda cambiar su condición de tratamiento o control. La manipulación de los criterios de elegibilidad puede adoptar numerosas formas. Por ejemplo, los encuestadores que recopilan los datos que se utilizan para calcular la puntuación de elegibilidad podrían cambiar una o dos respuestas de los encuestados; o puede que los encuestados mientan deliberadamente a los encuestadores si creen que con eso tendrán acceso al programa. Además, la manipulación de las puntuaciones puede agravarse a lo largo del tiempo, a medida que los encuestadores, los encuestados y los políticos comienzan a aprender las *reglas del juego*. En el ejemplo del **pei**, la manipulación en torno al límite se produciría si los CEO participante pudieran alterar los ingresos por innovaciones o si por ejemplo, dieran informes falsos financieros así como de sobre sus operaciones y facilidades. Por ejemplo, una **pbt** con insumos de x nivel reportara productividad para ser elegible en el programa, en el caso de que los beneficios previstos de la subvención a los insumos de innovación, merecieran la pena hacerlo. Una de las señales que delata la manipulación, se ilustra en la **Gráfica 7.6.**

Gráfica 7.6. Manipulación del índice de elegibilidad



Fuente: Elaboración propia.

El panel (a) muestra la distribución de las **pbt** según su índice de línea de base cuando no hay manipulación. La densidad de las **pbt** en torno al límite **(50)** es continua (o fluida). El panel (b) presenta una situación diferente: un número mayor de **pbt** parecen estar agrupados justo por debajo del límite, mientras que hay relativamente pocas **pbt** justo por encima del límite. Dado que no hay un motivo *a priori* para creer que debería haber un gran cambio en el número de **pbt** justo en torno al límite, la ocurrencia de ese cambio en la distribución en torno al límite es una prueba de que de alguna manera las **pbt** pueden estar manipulando sus puntuaciones para tener acceso al programa. Una segunda prueba de manipulación grafica el índice de elegibilidad en relación con la variable de resultado en la línea de base y verifica que no haya discontinuidad, o un *salto*, justo en torno a la línea del límite.

Ventajas y deventajas del enfoque RD

Khandker et al. (2017), señalan ventajas y desventajas de este enfoque, por ejemplo, las ventajas del método **RD,** son:

- Produce una estimación sin sesgo, del efecto de tratamiento en la discontinuidad.
- Que puede aprovechar en muchas ocasiones ventaja de una *regla conocida* para asignar el beneficio que es común en los diseños de política de innovación.
- Que un grupo de sujetos elegibles, no necesita ser excluido del tratamiento.

Sin embargo, las desventajas con **RD**, son:

- Que produce *efectos de tratamiento promedio local*, que no siempre son generalizables.

- Que el efecto se estima en la discontinuidad, entonces, generalmente, existen menos observaciones que en un experimento aleatorio con el mismo tamaño de la muestra.
- Que la especificación puede ser sensible a *la forma funcional*, incluyendo relaciones e interacciones no lineales.
- Una preocupación con el método **RD** es el *comportamiento* **(Ravallion, 2008).** Los administradores del programa es posible que no siempre conozcan con precisión, los *criterios de elegibilidad*; por lo tanto, las respuestas de comportamiento a la intervención del programa, pueden confundirse con *las reglas de focali*zación reales. Los datos recopilados antes de la intervención del programa, en forma de línea de base, por ejemplo, pueden ayudar a aclarar diseño del programa y su correspondiente captación.
- Otra preocupación, es que el ejercicio ,se centra solo en sujetos o unidades situados alrededor del umbral ***s*.*** Si este grupo es materialmente interesante para el el evaluador, debe ser abordado; si los funcionarios del programa están interesados, por ejemplo, en Identificar los efectos de un programa alrededor de una frontera geográfica y determinar si el programa debe expandirse a través de las fronteras, la muestra limitada es posible no sea de gran preocupación.
- Se puede construir un ejemplo similar, sobre un programa de incentivos a la innovación, de alivio relacionado con **pbt** cuyo estado se acerca a la línea de la insolvencia. Si las reglas de elegibilidad no se cumplen o cambian con el tiempo, la validez del enfoque de discontinuidad también necesita ser examinado más cuidadosamente. Comprobaciones de robustez se pueden realizar para examinar la validez del diseño de discontinuidad, incluso con cambios en otras variables de control del punto de corte.
- Examinar el patrón de la elegibilidad, para determinar variables, también puede ser útil, ya sea, por ejemplo, el resultado promedio que exhibe saltos en valores de la variable que no sean el límite de elegibilidad, como así como cualquier *discontinuidad* en la densidad condicional de esta variable. Si los datos de control exhiben no linea-

lidades, por ejemplo, una pendiente más pronunciada que los datos de tratamiento, luego el término cuadrado, para la variable de selección, se pueda agregar al resultado en la ecuación de regresión. Las no linealidades en la forma funcional, también se pueden abordar hacendo interactuar la variable de selección con el punto de corte o, quizás, mediante el uso de una regresión más corta y más densa líneas para capturar comparaciones más estrechas.

Comparaciones *pipeline*

Las comparaciones de *pipeline* explotan la variación en el tiempo, de la implementación del programa, utilizando como un grupo de comparación elegible, observaciones no dirigidas que aún no han recibido el programa. Se pueden usar con diseños de discontinuidad si se asigna un tratamiento sobre la base de algunas características exógenas y posibles participantes (quizás para un programa relacionado) que están esperando la intervención. Tal enfoque podría ser utilizado en el contexto de un programa en espera de expansión presupuestaria, por ejemplo, donde los sujetos en espera de tratamiento pueden usarse como un grupo de comparación. En esta situación, lo mismo se utilizaría el enfoque **RD**, pero con un subconjunto agregado (dinámico) de *no participantes*. Otro ejemplo podría ser donde una inversión local, como una capacitación en innovación en modelos de negocios textil, como fuente de mejoras adicionales del mercado, de modo que los sujetos alrededor de es tipo de giros, puedan beneficiarse de las futuras intervenciones; la variación en la exposición potencial en función de los insumos relacionados, podría explotarse como fuente de identificación **(Ravallion, 2008).**

Limitaciones e interpretaciones

Gertler et al. (2017) afirman que el diseño de regresión discontinua proporciona estimaciones del **LATE** en torno al límite de elegibilidad en el punto en que las unidades de tratamiento y comparación son más similares. Cuanto más se acerque uno a la puntuación límite, más similares serán las unidades a cada lado del umbral. De hecho, si uno se aproxima mucho a la puntuación límite, las unidades a ambos lados del umbral serán tan similares que su comparación será tan buena como si los grupos de tratamiento y de comparación se hubieran elegido mediante la asignación aleatoria del tratamiento.

Dado que el método de **RD** estima el impacto del programa en torno a la puntuación límite, o localmente, la estimación no se puede necesariamente generalizar a unidades cuyas puntuaciones se alejan más del umbral, es decir, donde los individuos elegibles y no elegibles quizá no sean tan similares. El hecho de que el método de **RD** no pueda proporcionar una estimación de un efecto de tratamiento promedio para todos los participantes del programa puede verse a la vez como una fortaleza y como una limitación, lo cual depende de la pregunta de la evaluación de interés.

Si la evaluación pretende responder la pregunta ¿el programa debería existir o no?, el efecto de tratamiento promedio para toda la población elegible puede ser el parámetro más relevante, y es evidente que el **RD** no será del todo perfecto. Sin embargo, si la cuestión de interés para las políticas es ¿el programa debería suspenderse o ampliarse?, es decir, para los beneficiarios (potenciales) justo en las inmediaciones del límite, la **RD** produce precisamente la estimación local de interés para sustentar esta importante decisión de política.

Como ya se señaló, pueden surgir otras complicaciones cuando el cumplimiento en cualquiera de los dos lados del límite es imperfecto. Este **RD** *difuso* se produce cuando las unidades que no son elegibles debido a su puntuación en el índice consiguen tener acceso al programa, o cuando las unidades elegibles según la puntuación del índice deciden no participar en el programa. En este caso, se puede utilizar una metodología de variable

instrumental similar a la definida en el capítulo 6, a saber, la localización de las unidades por encima o por debajo de la puntuación límite se usará como variable instrumental para la participación observada en el programa. Como sucedía en los ejemplos del capítulo 6, esto tiene un inconveniente: solo se puede estimar el impacto de aquellas unidades que son *sensibles* al criterio de elegibilidad, esto es si se trata del tipo Inscrito si es elegible, pero no si se trata del tipo *Siempres o Nuncas*.

El hecho de que el método de **RD** estime el impacto solo en las inmediaciones de las puntuaciones límite también genera dificultades en términos de la potencia estadística del análisis. En ocasiones, solo se emplea en el análisis un conjunto limitado de observaciones que se sitúan cerca de la puntuación límite, con lo cual el número de observaciones en el análisis de **RD** se reduce, en comparación con los métodos que analizan todas las unidades en los *grupos de tratamiento y comparación*. Para obtener una *potencia estadística* suficiente al aplicar el **RD**, habrá que utilizar un ancho de banda en torno a la puntuación límite que incluya un número suficiente de observaciones. En la práctica, se debería intentar utilizar un ancho de banda lo más amplio posible, a la vez que se conserva el equilibrio en las características observadas de la población por encima y por debajo de la puntuación límite.

Posteriormente, se puede aplicar la estimación varias veces usando diferentes anchos de banda para verificar si las estimaciones son sensibles al ancho de banda utilizado. Es necesario formular otra advertencia al utilizar el método de **RD**, a saber, la especificación puede ser sensible a la forma funcional que se emplea para modelar la relación entre la puntuación de elegibilidad y el resultado de interés. En los ejemplos presentados en este capítulo, se da por sentado que la relación entre el índice de elegibilidad y el resultado es lineal. En realidad, la relación podría ser más compleja, e incluir relaciones no lineales e interacciones entre variables. Si uno no se da cuenta de estas relaciones complejas en la estimación, se las puede confundir con una discontinuidad, lo que llevaría a una interpretación incorrecta de la estimación de impacto con **RD**. En la práctica, se puede estimar

el impacto del programa utilizando diversas formas funcionales (lineales, cuadráticas, cúbicas, cuárticas, y otras similares) para evaluar si, de hecho, las estimaciones de impacto son sensibles a la forma funcional.

Por último, como se señala más arriba, hay unas cuantas condiciones importantes para la regla de elegibilidad y el umbral **(Gertler et al., 2017):**

- En primer lugar, deben ser únicos del programa de interés. Por ejemplo, puede utilizarse un índice de pobreza que establezca un ranking de hogares o individuos para focalizar una diversidad de programas sociales para los pobres. En este caso, no será posible aislar el impacto de un solo programa de lucha específica contra la pobreza de todos los demás programas que utilizan los mismos criterios de focalización.
- En segundo lugar, la regla de elegibilidad y el umbral deberían ser resistentes a la manipulación de los encuestadores, los beneficiarios potenciales, los administradores de los programas o los políticos. La manipulación del índice de elegibilidad crea una discontinuidad en el índice que socava la condición básica para que el método funcione, a saber, que el índice de elegibilidad debería ser continuo en torno al umbral.

Incluso con estas limitaciones, **RD** es un poderoso método de evaluación de impacto para generar estimaciones no sesgadas del impacto de un programa en la cercanía del límite de elegibilidad. **RD** aprovecha las reglas de asignación del programa, a partir de índices de elegibilidad continuos, que ya son habituales en numerosos programas sociales. Cuando se aplican las reglas de focalización basadas en el índice, no es necesario excluir un grupo de sujetos como las **pbt** elegibles como beneficiarios del tratamiento a los fines de la evaluación, porque se puede utilizar el diseño de regresión discontinua como alternativa.

Lista de verificación: diseño de regresión discontinua

RD requiere que el índice de elegibilidad sea continuo en torno a la puntuación límite, y que las unidades sean similares en las cercanías por encima o por debajo de la puntuación límite, por lo que debe **(Gertler et al., 2017):**

- Es continuo el índice en torno la puntuación límite en el momento de la línea de base?
- Hay alguna evidencia de falta de cumplimiento de la regla que determine la elegibilidad para el tratamiento? Compruébese que todas las unidades elegibles y ninguna unidad no elegible han recibido el tratamiento. Si se encuentra falta de cumplimiento, habrá que combinar **RD** con un enfoque de variable instrumental para corregir esta *discontinuidad difusa*.
- ¿Hay alguna evidencia de que las puntuaciones del índice puedan haber sido manipuladas con el fin de influir en quien tenía derecho a beneficiarse del programa? Compruébese si la distribución de la puntuación del índice es fluida en el punto límite. Si se halla evidencia de una concentración de puntuaciones ya sea por encima o por debajo del punto límite, puede que esto sea una señal de manipulación.
- ¿El umbral corresponde a un único programa que se está evaluando o está siendo usado por otros programas también?

Ejemplo utilizando STATA

Cuando el tratamiento se asigna exclusivamente sobre la base de un valor de corte, entonces el diseño de la regresión de discontinuidad (RD) es una alternativa adecuada, a los experimentos aleatorios u otros diseños cuasi-experimentales. A diferencia del diseño aleatorio, un grupo elegible no necesita ser excluido del tratamiento, solo por el bien de la evaluación de impacto. En nuestro ejercicio, la evaluación de impacto puede implementarse con diseño de **RD**, utilizando los datos del **pei** para **pbt** porque la

participación en los programas de microcrédito están determinados oficialmente por la cantidad de activos, recursos, técnicos, financieros y humanos que se dediquen a la innovación; esto configura a una **pbt** elegible para participar solo si tiene menos de **50** decimales de dichos activos, como condición hipotética. Por lo tanto, El punto de corte de **50** decimales en los activos dedicado s a la innovación por **pbt**, es el criterio de diseño de **RD.**

Uso de RD en la estimación de impacto

La evaluación de impacto bajo el enfoque **RD**, se basa en la idea de que la muestra en el vecindario del *punto de corte* (arriba y abajo) representa características de diseño aleatorizado, porque los **pbt** en los grupos de tratamiento y control son muy similares en su características y varían solo en su *estado de tratamiento*. Entonces, una diferencia en la media de resultados de los grupos, tratados y de control restringidos a la vecindad del punto de corte, es decir, local a la discontinuidad, da el impacto de la intervención. **RD** tiene dos versiones:

- El llamado de *discontinuidad aguda* (*sharp discontinuity*), en el que el punto de corte se establece de manera determinista, el estado del tratamiento. Es decir, todos los elegibles reciben el tratamiento y nadie no elegible lo recibe.
- El otro tipo es llamado, de *discontinuidad difusa* (*fuzzy discontinuity*), el estado de tratamiento, no salta abruptamente de cero a uno a medida que los **pbt** se vuelven elegibles, desde inelegibles. Este escenario es más realista, particularmente en este caso, porque algunos **pbt** elegibles deciden (por una razón u otra) no participar en el microcrédito, mientras que algunos **pbt** no elegibles, participan. En un buen diseño **RD**, tanto los no participantes elegibles y los participantes inelegibles siguen siendo bajos. El impacto de la participación de acceso al microcrédito, mediante el diseño de **RD**, se puede dar mediante la siguiente expresión:

$$I = (y^{+} - y^{-}) / (s^{+} - s^{-}) \ldots\ldots\ldots\ldots\ldots (7.5)$$

Donde:

$\boldsymbol{y^{+}}$ es el resultado medio para los *participantes* de acceso al microcrédito cuyos activos y recursos técnicos, humanos y financieros dedicados a la innovación, están alrededor de **50** decimales, como referencia hipotética.
$\boldsymbol{y^{-}}$ es el resultado medio para los *no participantes* de acceso al microcrédito cuyos activos y recursos técnicos, humanos y financieros dedicados a la innovación, están alrededor de **50** decimales, como referencia hipotética.
$\boldsymbol{s^{+}}$ es el estado de tratamiento medio para **pbt** *elegibles* de acceso al microcrédito cuyos activos y recursos técnicos, humanos y financieros dedicados a la innovación, están alrededor de **50** decimales, como referencia hipotética.
$\boldsymbol{s^{-}}$ es el estado medio de tratamiento para **pbt** *no elegibles* de acceso al microcrédito cuyos activos y recursos técnicos, humanos y financieros dedicados a la innovación, están alrededor de **50** decimales.

En la *discontinuidad aguda* $\boldsymbol{s^{+} = 1}$ y $\boldsymbol{s^{-} = 0}$ y la diferencia en los resultados medios de participantes y no participantes da el impacto. En realidad, en lugar de calcular directamente los medios de resultado y tratamiento, se estima sus valores de regresiones locales lineales (o kernel) que se implementan en ambos lados del punto de corte. Luego, estos valores se conectan a la **ecuación 7.5** para obtener los impactos estimados.

Implementación de la discontinuidad aguda (*sharp discontinuity*)

Los datos de los archivos **pei_01.dta** o **pei_08.dta** no satisfacen las condiciones para cumplir el diseño de discontinuidad, porque la participación en el programa no es determinista basada en el punto de corte de los activos y recursos técnicos, humanos y financieros dedicados a la innovación, están alrededor de **50** decimales, como referencia hipotética. En otras palabras, algunos **pbt** *elegibles* (activos y recursos <**50** decimales) *no participan*, y **pbt** no elegibles (activos y recursos > = **50** decimales) participan. Por lo tanto, para demostrar una discontinuidad aguda, **pei_08.dta** son ajustados para abandonar a estos estos dos tipos de **pbt**:

- Teclear:

  ```
  use "/Users/DCA/Desktop/STATA log files/pei_08.dta"
  drop if (smefac<50 & (dmmfd==0|dfmfd==0))|(smefac>=50 & (dmmfd==1|dfmfd==1))
  ```

- El siguiente paso, es ejecutar la regresión lineal para el resultado (gasto **pbt** per cápita) vs. los activos, recursos, técnicos, financieros y humanos que se dediquen a la innovación tanto para *elegibles* (participantes) como *no elegibles* (no participantes) hogares. Como un resultado de la operación previa de la caída de algunas **pbt**, las que sean elegibles, ahora son determinísticamente participantes y las no elegibles como determinísticamente no participantes. La estimación por *regresión polinómica*, permite que los resultados se almacenen, tanto para los participantes como para los no participantes.
- Posteriormente, se deben tomar medidas de esos resultados en el punto de corte. Debido a que el punto de corte es un solo valor (**50** decimales) hipotético, es mejor especificar un rango de valores de activos y recursos técnicos, financieros y humanos de la **pbt** dedicados a la innovación, y tomar la media de los resultados, como referencia de los **pbt** que estén dentro de ese rango. Ese rango se establece

de **45** a **50** decimales, para participantes y de **50** a **55**, para no participantes. Con las medias de los resultados calculados, su diferencia se puede tomar para obtener los impactos estimados en la participación de las **pbt** de acceso al microcrédito basados en sus gastos per cápita en el vecindario del punto de corte. Todo este proceso se codifica de la siguiente manera dentro del programa **STATA**, llamado *rd_sharp*:

```
clear
use “/Users/DCA/Desktop/STATA log files/pei_08.dta”
drop if (smefac<50 & (dmmfd==0|dfmfd==0))|(smefac>=50 & (dmmfd==1|dfmfd==1))
prog rd_sharp, rclass
*version 14.1
args outcome
confirm var `outcome’
tempname outrd1 outrd0 outcome1 outcome0
lpoly `outcome’ lnsmefac if smefac<50, gen(`outrd1’)
at(ln smefac) nogr tri w(3) d(1)
lpoly `outcome’ lnsmefac if smefac>=50, gen(`outrd0’)
at(lnsmefac) nogr tri w(3) d(1)
sum `outrd1’ if smefac>=45 & smefac<50, meanonly
scalar `outcome1’=r(mean)
sum `outrd0’ if smefac>=50 & smefac<55, meanonly
scalar `outcome0’=r(mean)
return scalar diff_outcome=`outcome1’-`outcome0’;
end
```

Aunque los cálculos estimados de impacto, se pueden calcular de esta manera, este proceso no nos reporta el error estándar, el cual, se utiliza para calcular las estadísticas ***t***. El error estándar puede calcularse ejecutando el programa *Bootstrapping*, el cual ejecuta un comando (o un conjunto de comandos) repetidamente y al azar con reemplazo, tomando en cuenta las observaciones, almacenando los resultados por cada ejecución y poste-

riormente, calcular el error estándar de las estimaciones guardadas. No es necesario que cada comando (*bootstrapped*) se ejecute por separado. En su lugar, un programa que incluya todos los comandos necesarios, es muy conveniente para ejecutarse con *bootstrap*. La programación, también permite ejecutar el mismo programa usando diferentes parámetros. Por ejemplo, observe las diferentes opciones del comando *lpoly* en el programa *rd_sharp*, que realiza la *regresión lineal polinóimica* de la variable de resultado genérico vs. el registro de activos y recursos técnicos, financieros y humanos que dedica la **pbt** a la innovación, para **pbt** participantes y no participantes. Así, se consideran:

gen() que almacena el resultado de la estimación.
at() que especifica una variable que contiene, los valores en los cuales, una regresión suave de kernel que debe ser evaluada.
tri que especifica que el tipo de regresion lineal kernel, es triangular.
w que especifica la banda media del kernel, el ancho de banda de la ventana suavizada alrededor de cada punto.
nogr suprime gráficos para cada ancho de banda.
d() especifica el grado del polinomio a ser usado en la suavidad (1 implica regresión lineal).

En la regresión lineal polinómica, diferentes anchos de banda pueden producir diferentes estimaciones, por lo que se recomienda realizar pruebas con más de un ancho de banda. La elección del kernel es menos importante, aunque probar diferentes tipos puede ayudar a verificar la solidez de las estimaciones. Una observación importante para hacer aquí, es que el programa *rd_sharp* no tiene ningún parámetro para indicar la participación en el programa de acceso al microcrédito. Esto es, porque la participación al **pei** se ha hecho determinista por el registro de los activos y recursos técnicos, financieros y humanos que dedica la **pbt** a la innovación. Los siguientes comandos relacionan al *bootstrapping,* el cual, se realiza ejecutando el comando **STATA** *bootstrap*, que debe ser seguido por el comando y luego la expresión estadística o expresión a estimar.

El comando *bootstrap*, ejecuta el programa *rd_sharp* definido previamente, con el argumento *lexptot*, que reemplaza el argumento genérico *outcome* con *lexptot* (*log of gasto anual per capita*). En consecuencia, *lexptot* se ejecuta contra *lnsmefac* (registro de activos y recursos técnicos, financieros y humanos que dedica la **pbt** a la innovación) utilizando regresiones lineales polinómicas. Al final de la ejecución, el programa *rd_sharp* calcula la diferencia de las medias de *lexptot* (impacto estimado), donde el comando *bootstrap* almacena sus datos en una variable llamada *impact_sharp*. Finalmente, *bootstrap* ejecuta el *rd_sharp* programa **100** veces. Así:

- Teclear:

  set seed 12345
  bootstrap (rd_sharp lexptot) impact_sharp=r(diff_outcome),
  reps(100) nowarn

- La salida del comando *bootstrap*, muestra la participación de acceso al microcrédito en el **pei** tiene un impacto negativo en el gasto per cápita (**−12.6 %**) y el error estándar es **0.1116.**

Ver Tabla 7.1.

Tabla 7.1. Tabulación resultado de aplicar el método RD aguda (sharp continuity)

```
command:      rd_sharp lexptot
statistic:    impact_s~p = r(diff_outcome)

Bootstrap statistics                            Number of obs    =       243
                                                Replications     =       100

------------------------------------------------------------------------------
Variable     | Reps  Observed    Bias     Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
impact_sharp |   92  -.1264224  .0023491   .1116639  -.3482292 .0953843   (N)
             |                                       -.3132059 .0937947   (P)
             |                                       -.3132059  .125849  (BC)
------------------------------------------------------------------------------

Note: N   = normal
      P   = percentile
      BC  = bias-corrected
```

Fuente: STATA con datos propios.

- Los siguientes comandos crean una ***t*** estadística de los impactos estimados que los despliega:

  gen t_impact_sharp=_b[impact_sharp]/_se[impact_sharp]
  sum t_impact_sharp

- Después de la ejecución de los comandos, se observa que el impacto estimados no es sigificante **(*t* =-1.132). Ver Tabla 7.2**

Tabla 7.2. Tabulación resultados *t* stadística del método RD aguda

Variable	Obs	Mean	Std. Dev.	Min	Max
t_impact_s~p	243	-1.132169	0	-1.132169	-1.132169

Fuente: STATA con datos propios.

Implementación de la discontinuidad difusa (*Fuzzy Discontinuity*)

A diferencia de la implementación de la regresión por *discontinuidad aguda*, la implementación de regresión por *discontinuidad difusa* no requiere abandonar las observaciones para la no participación de hogares elegibles o participación de hogares no elegibles. El programa para estimar los impactos de la *discontinuidad difusa* es muy similar al utilizado para la *discontinuidad aguda*. Aquí se incluyen tambié *regresiones polinómicas* para su tratamiento, además de las de los resultados. El impacto estimado se calcula utilizando la fórmula especificada en la **ecuación 7.5**. El programa para calcular *la discontinuidad difusa,* es:

- Teclear:

  clear
  use "/Users/DCA/Desktop/STATA log files/pei_08.dta"

```
drop if (smefac<50 & (dmmfd==0|dfmfd==0))|(smefac>=50 &
(dmmfd==1|dfmfd==1))
prog rd_fuzzy, rclass;
*version 14.1
args treatment outcome
confirm var `treatment’
confirm var `outcome’
tempname treatrd1 treatrd0 outrd1 outrd0 treat1 treat0 outcome1
outcome0
lpoly `treatment’ lnsmefac if smefac<50, gen(`treatrd1’)
at(lnsmefac) nogr tri w(3) d(1)
locpoly `treatment’ lnsmefac if smefac>=50, gen(`treatrd0’)
at(lnsmefac) nogr tri w(3) d(1)
locpoly `outcome’ lnsmefac if smefac<50, gen(`outrd1’)
at(lnsmefac) nogr tri w(3) d(1)
lpoly `outcome’ lnsmefac if smefac>=50, gen(`outrd0’)
at(lnsmefac) nogr tri w(3) d(1)
sum `treatrd1’ if smefac>=45 & smefac<=55, meanonly
scalar `treat1’=r(mean)
sum `treatrd0’ if smefac>=45 & smefac<=55, meanonly
scalar `treat0’=r(mean)
```

- El programa *rd_fuzzy*, a diferencia de *rd_sharp*, toma dos argumentos, uno para el *tratamiento* y otro para el *resultado*. Por lo tanto, para estimar los impactos de la participación de CEOs femeninas de las **pbt** en el **pei** de acceso a microcrédito sobre el gasto per cápita de las **pbt**, el comando *bootstrap* ejecuta el programa *rd_fuzzy* con dos argumentos: *dfmfd* (participación de microcréditos de las CEO femeninas) y *lexptot* (gasto anual per cápita). Por lo tanto, se teclea:

```
set seed 123;
bootstrap (rd_fuzzy dfmfd lexptot) impact_fuzzy_f=r(impact),
reps(100) nowarn
```

- La salida del comando *bootstrap* muestra que el signo de impacto estimado sigue siendo negativo (**ver Tabla 7.3**):

Tabla 7.3. Resultado de ejecutar bootstrap en RD difusa

```
command:      rd_fuzzy dfmfd lexptot
statistic:    impact_f~f = r(impact)

Bootstrap statistics                          Number of obs    =      1129
                                              Replications     =       100

------------------------------------------------------------------------------
Variable     | Reps  Observed    Bias    Std. Err. [95% Conf. Interval]
-------------+----------------------------------------------------------------
impact_fuz~f | 100  -1.702198  1.92124   3.571683  -8.789193   5.384796  (N)
             |                                    -10.52238   9.24404   (P)
             |                                    -13.93708  -.0473376 (BC)
------------------------------------------------------------------------------

Note: N  = normal
      P  = percentile
      BC = bias-corrected
```

Fuente: STATA con datos propios.

- Los siguientes comandos crean el resultado de las estadisticas ***t*** del impacto estimado:

  gen t_impact_fuzzy_f=_b[impact_fuzzy_f]/_se[impact_fuzzy_f]
  sum t_impact_fuzzy_f

- Después de ejecutar estos comandos, se observa que el impacto es significante **(*t* = -0.477) .Ver Tabla 7.4**

Tabla 7.2. Tabulación resultados t stadística del método RD difusa

Variable	Obs	Mean	Std. Dev.	Min	Max
t_impact_f~f	1129	-.4765815	0	-.4765815	-.4765815

CAPÍTULO 8.

Elección del Método a Utilizar

Aunque no esxisten reglas tácitas para decir qué técnica es mejor que otra, lo que es factor clave que identifica el impacto causal de un programa consiste en **(Gertler et al., 2017):**

- Encontrar un grupo de comparación válido para estimar el *contrafactual*.
- Responder a la pregunta de interés de la política de innovación de las autoridades que rigen las condiciones del proyecto.

Al momento, se conocen las condiciones de los métodos de: **aleatorización,** propensión de coincidencia de puntaje **(PSM)**, doble diferencia **(DD)**, variable instrumental **(IV)**, Regresión discontinua **(RD)**. Por lo tanto, para determinar qué método utilizar se recomienda:

- Considerar que las reglas operativas del programa, son la base que esclarece a los *grupos de comparación* que facilita hallar el método más adecuado dentro del ámbito de las políticas de implementación de las innovaciones. A mayor definición de los alcances de las reglas operativas de un programa, mayor auxilio para determinar el método más adecuado que evalúe a ese programa en particular.
- Los métodos tienen diferentes condiciones, algunas más estrictas que otras, que deben reunir los datos con diversos supuestos básicos para estimar con precisión los cambios en los resultados debido a la intervención.
- Son preferibles los métodos con supuestos más débiles y con la menor cantidad de requisitos de datos a cubrir, respecto al contexto de las reglas operativas.

- Es vital analizar cómo elegir la unidad de intervención, desde el nivel individual, pymes, sectores, etc. hasta órdenes superiores como las comunidades, clusters, distritos, regiones, estados, países. Es muy útil que la unidad de intervención factible sea las más pequeña dentro de las limitaciones operativas.

Lo interesante es que los grupos de comparación usen las reglas operativas del programa para ser determinadas, siempre y cuando, las reglas de operación se encuentren bien definidas. Esto significa, que existe una orientación de un solo sentido: *las reglas de operación de un programa deben regir en el métofo de evaluación, no a la inversa*. Es muy mala señal, que el método de evaluación cambie drásticamente elementos clave de las reglas de operación de un programa. Una regla de operación clara, debe:

- Identificar los sujetos elegibles y determinar como se seleccionan.
- Los grupos de comparación, provienen o:
  - De los sujetos elegibles pero que en determinados momentos, no se pueden incluir,ndebido a varias causas, como recursos escasos, sobredemanda, etc. o
  - De los sujetos que se ubican próximos al umbral de elegibilidad para participar en el programa.

Así, una evaluación de impacto correctamente diseñada, define sus grupos de comparación válidos cuando se presentan las reglas de operación de los participantes, muestran reglas equitativas (en función a indicadores acordados donde todos tienen las mismas posibilidades de acceso a los beneficios), las reglas de divulgación al programa son transparentes (existe seguimiento observable, claro y preciso) y sujetas a auditoría por parte de los administradores en la introducción de innovaciones. Se crean problemas si estas condiciones no se cumplen.

La importancia de determinar los grupos de comparación

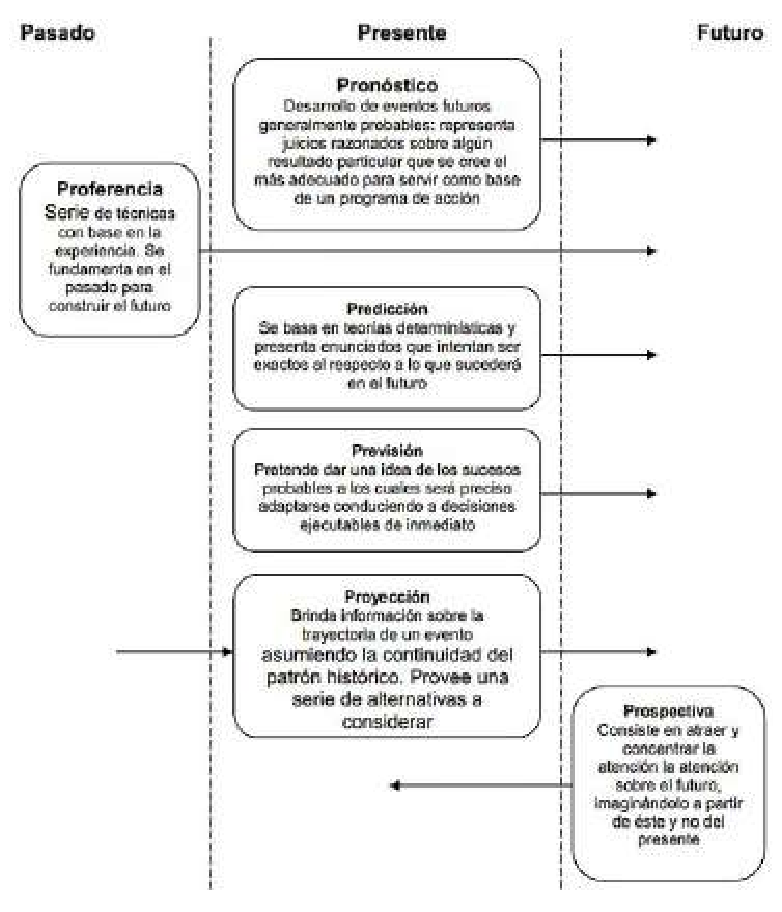
Es posible establecer un plan de ruta si se responden tres preguntas operativas fundamentales vs. los recursos disponibles del programa:

- *Los recursos disponibles* o ¿Existen suficientes recursos, por parte del programa, para una implementación a escala y atender a los participantes elegibles? Responder esta pregunta es la que impone las condiciones de geografía, de ingreso económico, de tipo de pyme participante, etc.
- *Los criterios de elegibilidad* o ¿quién es elegible para recibir los benefcios del programa?, ¿existe un umbral de elegibilidad o está disponible el acceso a todos sin restricciones? Responder estas preguntas, permite definir más claramente al grupo beneficiario
- *Calendario de implementación* o ¿los beneficiarios participantes potenciales se inscriben al programa al mismo tiempo o son asignados por etapas? Responder estas preguntas, permite optimizar los recursos finitos del programa.

Prospectiva y cómo crear un grupo de comparación

El estudio del futuro, en la implementación de todo proyecto de innovación con impacto social, juega un papel importante en los actores ya sea de tratamiento o de control. Tomando de referencia al **DRALE (2020)**, la palabra futuro proviene del latín *futurus*, significando: *lo que está por llegar o suceder*; así **Medina** y **Ortegón (2006, en Mejía-Trejo, 2011**) informan sobre 6 enfoques para su estudio (los 4 primeros, parten del presente al futuro; el 5º del pasado al futuro y el último, del futuro al presente: Existen al momento, 6 enfoques para su estudio (los 4 primeros, parten del presente al futuro; el 5º del pasado al futuro y el último, del futuro al presente. Ver **Figura 8.1.**

Figura 8.1. Diferentes enfoques sobre los estudios del futuro



Fuente: Medina y Ortegón (2006, en Mejía-Trejo, 2011).

Así, la técnica considerada como más lograble para crear los grupos de comparación, es la prospectiva ya que permite conocer no uno, sino varios futuros, así como estudiarlos, evaluarlos y seleccionar el más conveniente dentro de lo factible. La prospectiva contempla el futuro cercano como etapa para la construcción de un futuro lejano más conveniente. Se trabajan las acciones del presente en función del *futuro deseado, probable y posible*, sin por ello desaprovechar el pasado y un presente conocidos con

relativa suficiencia. Además de diseñar el futuro deseable y hacerlo posible, aporta una serie de elementos al proceso de toma de decisiones y, lo que es más importante, facilita el ponerse de acuerdo en forma concertada hacia el mejor futuro posible y que, después de ello, se actúe en concordancia y se logre. Así, requiere y genera actitudes tanto creativas como realistas y constructivas hacia el devenir. Ver **Tabla 8.1**

Tabla 8.1. Diferentes enfoques del estudio del futuro

Vía	Futuros		
	Deseable	Posible	Probable
Prospectiva	X	X	X
Proferencia	-	-	X
Pronóstico exploratorio	-	-	X
Pronóstico normativo	X	-	-
Predicción	-	-	X
Previsión	-	-	X
Proyección	-	-	X

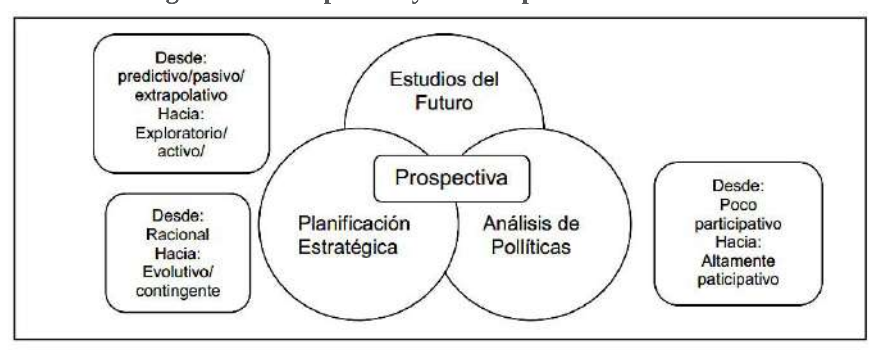
Fuente: Medina y Ortegón (2006 en Mejía-Trejo, 2011).

La prospectiva, según el **DRALE (2020)** proviene del latín *prospectio-onis*, que significa: *exploración de posibilidades futuras basada en indicios presentes*, lo cual, de acuerdo a lo analizado, podría parecer un error. Lo debemos concebir como *el esfuerzo de hacer probable el futuro más deseable; tiene trayectoria del porvenir hacia el presente;* se construye el mañana a través de *futuribles (futuro posible)* y *futurables (futuro deseable); su propósito es preparar el camino para el futuro, adoptándolo como objetivo (deseable y posible)* **(Medina y Ortegón, 2006, en Mejía-Trejo, 2011).** Es conocida también como *prospective* en lengua francesa, *foresight* en lengua inglesa o *prospeçcao* en lengua portuguesa, se define así como un proceso de anticipación y exploración de la opinión experta

proveniente de redes de personas e instituciones del gobierno, la empresa y las universidades, en forma estructurada, interactiva y participativa, coordinada y sinérgica, para construir visiones estratégicas de la ciencia y la tecnología y su papel en la competitividad y el desarrollo de un país, territorio, sector económico, empresa o institución pública.

La prospectiva es una disciplina para el análisis de sistemas sociales, que permite conocer mejor la situación presente, identificar tendencias futuras y analizar el impacto del desarrollo científico y tecnológico en la sociedad. Con ello se facilita el encuentro entre la oferta científica y tecnológica con las necesidades presentes y futuras de los mercados y de la sociedad. A la par, los ejercicios movilizan a los diferentes actores sociales para generar visiones compartidas de futuro, orientar políticas de largo plazo y tomar decisiones estratégicas en el presente, dadas las condiciones y las posibilidades locales, nacionales y globales. La prospectiva se constituye así en un campo en plena evolución, de intersección entre los estudios del futuro, el análisis de políticas públicas para implementar innovaciones y la planificación estratégica. Fundamentalmente, busca aclarar las prioridades gubernamentales y de la región, sector o cadena productiva bajo estudio. Pero su propósito más amplio es promover un gran cambio cultural, una mejor comunicación, una interacción más fuerte y una más grande comprensión mutua entre los actores sociales para pensar su futuro y tomar decisiones desde el presente **(Gavigan, et al. 2002, en Mejía-Trejo, 2011).** Ver **Figura 8.2.**

Figura 8.2. Prospectiva y sus campos de intersección



Fuente: Gavigan et al., (2002), en Mejía-Trejo (2011).

Al diseñar evaluaciones de *impacto prospectivas* considerando los escenarios *posibles, probables y deseables*, hacen que la respuesta a las tres preguntas operativas, determine en gran parte el método de evaluación de impacto más adecuado, para un cierto programa. **Ver Tabla 8.2.** que muestra, los posibles grupos de comparación en relación con las reglas operativas específicas del programa y las tres preguntas operativas fundamentales relacionadas con los *recursos disponibles, las reglas de elegibilidad y el calendario de implementación*.

Las columnas se dividen en función de si el programa tiene o no recursos suficientes para eventualmente cubrir a todos los beneficiarios elegibles potenciales (*recursos disponibles*) y, además, se subdividen en programas que tienen un ranking y un límite de elegibilidad continuos y aquellos que no los tienen (*criterios de elegibilidad*). Las filas se dividen en implementación en fases o implementación inmediata del programa (*calendario de implementación*). En cada celda se recogen las fuentes potenciales de grupos de comparación válidos, junto con el capítulo relacionado en que se trata en la segunda parte. Cada celda tiene un rótulo con un índice: la letra inicial señala la columna en el cuadro **(A, B)** y el número que sigue indica la columna **(1-4).** Por ejemplo, la celda **A1** se refiere a la celda de la primera fila y la primera columna del cuadro. Así, la celda **A1** identifica los métodos de evaluación más adecuados para los programas que tienen

recursos limitados, que tienen criterios de elegibilidad y se desarrollan en fases.

Tabla 8.2. Métodos de evaluación vs. reglas operativas de un programa

Calendario de Implementación		**Exceso de demanda del programa (recursos limitados)**		**No exceso de demanda del programa (recursos insuficientes)**	
	Citerios de elegibilidad	**(1) ICyUE**	**(2) SICyUE**	**(3) ICyUE**	**(4) SICyUE**
	(A) Implementación en fases	**Celda A1** AA RD	**Celda A2** AA IVPA DD DD PSM	**Celda A3** AAEF RD	**Celda A4** AAEF IVPA con participactón temprana DD DD PSM
	(B) Implementación inmediata	**Celda B1** AA RD	**Celda B2** AA IVPA DD	**Celda B3** RD	**Celda B4** Si la participación no es plena: IVPA DD DD PSM

Notas:
ICyUE. Índice Contínuo y Umbral de Elegibilidad
SICyUE.Sin Índice Contínuo y Umbral de Elegibilidad
AA. Asignación Aleatoria. (Capítulo 3)
AAEF. Asignación Aleatoria en Fases. (Capítulo 3)
DD. Doble Diferencia. (Capítulo 5)
DD PSM. Diferencia en Diferencias con Coincidencia o Pareamiento. (Capítulo 4)
RD. Regresión Discontinua.(Capítulo 7)
IVPA. Variables Instrumentales con promoción aleatoria.(Capítulo 6)

Fuente: Gertler et al. (2017).

Hay casos a describir por celdas, por ejemplo, la mayoría de los programas:

- Debe implementarse en fases a lo largo del tiempo debido ya sea a limitaciones financieras o a problemas logísticos y administrativos. Este grupo o categoría cubre la primera fila del cuadro **(celdas A1, A2, A3 y A4).** En este caso, la regla operativa equitativa, transparente y sujeta a rendición de cuentas consiste en dar a todas las unidades elegibles la misma oportunidad de ser la primera, segunda, tercera, etc. en acceder al programa, lo que implica una implementación aleatoria del programa a lo largo del tiempo.
- En los casos en que los recursos son limitados, es decir, en los que nunca habrá suficientes recursos para alcanzar la plena implementación **(celdas A1 y A2, y B1 y B2),** puede producirse muy rápidamente un exceso de demanda de aquellos recursos. Un sorteo para decidir quién entra en el programa puede ser un enfoque viable para decidir a quién asignar beneficios entre unidades igualmente elegibles. En este caso, cada unidad elegible tiene la misma oportunidad de beneficiarse del programa. Un sorteo es un ejemplo de regla operativa equitativa, transparente y sujeta a rendición de cuentas para asignar los beneficios del programa entre las unidades elegibles.
- Otro tipo de programas comprenden a los que se implementan a lo largo del tiempo y para los que los administradores pueden clasificar los beneficiarios potenciales en función de la necesidad **(celdas A1 y A3).** Si los criterios utilizados para priorizar a los beneficiarios son cuantitativos, están disponibles y tienen un umbral de elegibilidad, el programa puede usar un diseño de regresión discontinua.
- La otra categoría amplia consiste en programas que tienen la capacidad administrativa para implementarse inmediatamente: es decir, las celdas en la fila inferior del cuadro. Cuando el programa tiene recursos limitados y no es capaz de clasificar a los beneficiarios **(celda B2),** podría utilizarse la asignación aleatoria basada en el exceso de demanda. Si el programa tiene suficientes recursos para

ampliarse y ningún criterio de elegibilidad (celda B4), la única solución es utilizar variables instrumentales (promoción aleatoria) bajo el supuesto de participación no plena en el programa. Si el programa puede clasificar a los beneficiarios y depende de criterios de elegibilidad, se puede recurrir a la regresión discontinua.

Beneficiarios y su priorización

Para encontrar grupos de comparación válidos, el tema fundamental de cómo se seleccionan los beneficiarios, guarda relación con las tres preguntas operativas clave. En ocasiones, los *grupos de comparación* se hallan entre las *poblaciones no elegibles*, y con mayor frecuencia entre las *poblaciones que son elegibles* pero que se incorporan al programa más tarde. La manera de priorizar entre los beneficiarios depende en parte de los objetivos del programa ¿Se trata de un programa de incentivos para la innovación de las **pbt** sin ganacias, un programa de incremento de la creatividad focalizado pymes tradicionales o un programa para diseñar productos innovadores disponible para todos? Para priorizar entre los beneficiarios sobre la base de la necesidad, el programa debe encontrar *un indicador que sea a la vez cuantificable y verificable*. En la práctica, la viabilidad de la priorización depende en gran parte de la capacidad del administrador de innovaciones, para medir y clasificar las necesidades. Si el administrador de innovaciones clasifica adecuadamente a los beneficiarios en función de sus necesidades relativas, es posible que esté éticamente obligado a implementar el programa de acuerdo con las necesidades. Sin embargo, clasificar en función de la necesidad, requiere no solo una medida cuantificable sino también la capacidad y los recursos para medir ese indicador para cada unidad que participa en el programa.

Algunos programas utilizan criterios de selección que, en principio, podrían usarse para clasificar necesidades relativas y determinar la elegibilidad. Por ejemplo, numerosos programas quieren llegar a las **pbt** o pymes tradicionales de escaspos recursos. Sin embargo, los indicadores de pobreza adecuados que clasifican a los **pbt** de manera fiable a menudo son difíciles

de medir y costosos de recopilar. La recopilación de datos de los ingresos o del consumo de todos los beneficiarios potenciales para clasificarlos según el nivel de pobreza es un proceso complejo y oneroso que, además, sería difícil de verificar. Algunos programas utilizan *proxy mean test* para estimar los niveles de pobreza. Se trata de índices de medidas observables sencillas como los activos y las características sociodemográfica. Los *proxy mean tests* ayudan a determinar, razonablemente bien si una **pbt** por ejemplo, se sitúa por encima o por debajo de un umbral, pero pueden ser *menos precisos* en una clasificación detallada de la situación socioeconómica o de las necesidades. Numerosos programas han decidido clasificar en un nivel superior de agregación, como el nivel de la comunidad, en lugar de enfrentarse al costo y a la complejidad de clasificar a los potenciales beneficiarios individuales. Determinar la asignación del programa a un nivel agregado tiene beneficios operativos evidentes, pero a menudo es difícil encontrar indicadores para producir una clasificación de las necesidades en un nivel más agregado.

Se tienen que usar otros criterios para decidir cómo secuenciar la implementación del programa, en los casos en que un programa no pueda asignar beneficios de manera fiable sobre la base de la necesidad, ya sea porque es demasiado caro y propenso a errores o porque no hay indicadores de clasificación cuantificables y verificables.

La equidad, es un criterio coherente con la buena gobernanza, esto es, dar a todos aquellos que son elegibles, la misma oportunidad de ser el primero en tener acceso, y asignar de forma aleatoria un lugar en la secuencia a los beneficiarios potenciales, será una regla equitativa. Dadas las dificultades para clasificar las necesidades, en la práctica, una regla de asignación al programa que suele usarse, es la *asignación aleatoria* de los beneficios del programa. El diseño de evaluación aleatoria es capaz de entregar validez interna razonable, si es bien implementada, y es posible que dependa de supuestos más débiles en comparación con los otros métodos.

Comparando a los métodos de evaluación de impacto

Tomando en cuenta las reglas operativas específicas del programa y una vez determinado qué método de evaluación de impacto es adecuado, el equipo de evaluación debe elegir el método que tiene el supuesto más débil y los menores requisitos de datos. La **Tabla 8.3** muestra una comparación de los métodos de evaluación de impacto alternativos en términos de los requisitos de datos para implementarlos y los supuestos fundamentales necesarios para interpretar sus resultados como impactos causales de la intervención, de forma que:

- Cada fila representa un método diferente. Las primeras dos columnas describen los métodos y las unidades en el grupo de comparación. Las dos últimas columnas recogen los supuestos necesarios para interpretar los resultados como causales, y los datos necesarios para implementar los métodos.

Tabla 8.3. Comparación de métodos de evaluación de impacto

Metodología	Descripción	¿Quiénes están en el grupo de comparación?	Supuestos clave	Datos requeridos
Asigación aleatoria	Las unidades elegibles se asignan de forma aleatoria a un grupo de tratamiento o de comparación. Cada unidad elegible tiene una probabilidad conocida de ser seleccionada. Tiende a generar estimaciones de impacto internamente válidas con los supuestos más débiles.	Las unidades elegibles se asignan aleatoriamente al grupo de comparación.	La aleatorización produce dos grupos estadísticamente idénticos con respecto a las características observables y no observables a lo largo del tiempo en ausencia de la intervención (en la línea de base y a lo largo del seguimiento).	Datos de seguimiento de los resultados en los grupos de tratamiento y comparación; datos de línea de base y otras características para los grupos de tratamiento y comparación con el fin de verificar el equilibrio.

Variables instrumentales (concretamente la promoción aleatoria)	Un instrumento aleatorizado (como una campaña de promoción) induce cambios en la participación en el programa que se evalúa. El método utiliza el cambio en los resultados inducido por el cambio en las tasas de participación para estimar los impactos del programa.	Las unidades que cumplen con los requisitos para participar pero cuya participación se ve afectada por el instrumento (participarían si se exponen al instrumento pero no lo harían en caso contrario).	El instrumento afecta la participación en el programa, pero no afecta directamente los resultados (es decir, el instrumento influye en los resultados solo cambiando la probabilidad de participar en el programa).	Datos de seguimiento de los resultados de todas las unidades; datos sobre la participación efectiva en el programa; datos de los resultados de línea de base y otras características.
Diseño de regresión discontinua	Las unidades se clasifican a partir de criterios cuantitativos específicos y continuos, como un Índice de pobreza. Un umbral determina si una unidad es elegible para participar en un programa. Los resultados de los participantes en una parte del umbral se comparan con los resultados de los no participantes al otro lado del umbral.	Las unidades situadas cerca del umbral, pero que no son elegibles para recibir el programa.	Para identificar impactos no sesgados en el programa para la población cercana al umbral, las unidades que se encuentran inmediatamente por debajo e inmediatamente por encima del umbral son estadísticamente idénticas. Para identificar los impactos no sesgados en el programa para toda la población, la población cercana al umbral debe ser representativa de toda la población.	Datos de seguimiento de los resultados; Índice de clasificación y umbral de elegibilidad; datos sobre los resultados de línea de base y otras características.

Diferencias en diferencias	El cambio en el resultado a lo largo del tiempo en un grupo de no participan tes se utiliza para estimar cuál habria sido el cambio en los resultados de un grupo de participantes en ausencia de un programa.	Las unidades que no participaron en el programa (por cualquier motivo) y para las cuales se recopilaron datos antes y después del programa.	Si el programa no existia, los resultados de los grupos de participantes y no participantes habrian evolucionado paralelamente a lo largo del tiempo.	Datos de línea de base y de seguimiento de los resultados y otras características tanto para los participantes como para los no participantes.
Paramiento (en particular, pareamiento por puntajes de propensión)	Para cada participante del programa, el método busca la unidad *más similar* en el grupo de no participantes (el pareamiento más estrecho se basa en caracteristicas observables).	Para cada participante, la unidad no participante que, según las predicciones sobre la base de caracteristicas observables, tiene la misma probabilidad de haber participado en el programa.	No hay ninguna caracteristica que influya en la participación en el programa más allá de las características observables utilizadas para el pareamiento.	Seguimiento de los datos de los resultados de los participantes y no participantes; datos sobre la participación efectiva en el programa; caracteristicas de línea de base para llevar a cabo el pareamiento.

Fuente: Gertler et al. (2017):

- Todos los métodos parten de supuestos, es decir, para ser capaces de interpretar resultados como causales, se debe creer que son verdad ciertos hechos que no siempre se pueden verificar empíricamente. Para cada método, un supuesto clave es que la media del grupo de comparación de la que depende el método, sea una estimación válida del contrafactual.
- En cada método presentado, se han expuesto algunas consideraciones sobre cómo probar si un método es válido en un contexto particular. Algunos métodos dependen de supuestos más fuertes que otros.
- El método preferido es el que mejor se adecua al contexto operativo y el que requiere los supuestos más débiles y la menor cantidad de

datos. Estos criterios explican por qué los investigadores consideran la asignación aleatoria la regla de oro, y por qué a menudo es el método preferido.

- La asignación aleatoria, se adecúa a numerosos contextos operativos y tiende a generar estimaciones de impacto internamente válidas con los supuestos más débiles. Cuando se implementa de manera adecuada, genera comparabilidad entre los grupos de tratamiento y comparación en características observables y no observables. Además, la asignación aleatoria tiende a requerir muestras más pequeñas que las necesarias para implementar métodos *cuasi-experimentales*. Dado que la asignación aleatoria es relativamente intuitiva, el método también facilita la comunicación de resultados a los responsables de las políticas. Los experimentos sociales pueden generar estimaciones confiables de los impactos de programas nuevos y antiguos, pero tienen costos y beneficios que se deben ponderar antes de decidir si se eligen como estrategia de evaluación. En caso de adoptarse, esta herramienta requiere una implementación cuidadosa y una buena interpretación de los resultados. Dado que por diseño resuelven el problema del sesgo de selección, los experimentos aleatorios son considerados como punto de referencia para determinar el impacto de los programas o intervenciones. Entre sus ventajas se cuenta con que los resultados son muy transparentes, dado que, por la calidad de los datos generados, no es necesario utilizar herramientas econométricas sofisticadas. Por lo tanto ayudan, a la transparencia de las políticas públicas y a la rendición de cuentas por auditoría. Los experimentos sociales, a pesar de sus ventajas, sufren diversas limitaciones. Por un lado, pueden ser costosos y de difícil implementación. Esta desventaja se ha visto atenuada, pues las evaluaciones experimentales que se han llevado a cabo últimamente en los países en desarrollo son más modestas y a una escala mucho menor. Por tanto, se han vuelto una alternativa seria en las evaluaciones **(Duflo,et al., 2008)**. Segundo, la experimentación sufre problemas éticos al negarle a una porción de la población elegible los beneficios de la intervención. Sin embargo,

hay variantes del diseño puramente experimental que aminoran este problema, por ejemplo, aprovechar las restricciones logísticas para la ampliación de un programa. Además, dado el diseño de los experimentos y la imposibilidad de negar los beneficios a un grupo de control por largos periodos, con frecuencia es imposible estimar los impactos de largo plazo. Es posible que los efectos de largo plazo difieran de manera sustancial de los de corto plazo. Los experimentos también tienen problemas de validez externa que se pueden suavizar repitiendo el experimento en distintas regiones, experimentando con diferentes versiones del programa o usando la teoría económica para entender los canales a través de los cuales se generan los efectos. Es posible que los métodos cuasi-experimentales sean más adecuados en algunos contextos operativos, pero requieren más supuestos con el fin de que el grupo de comparación provea una estimación válida del contrafactual. **(Bernal y Peña, 2017).**

- El método **DD**, por ejemplo, depende del supuesto de que los cambios en los resultados en el grupo de comparación proporcionen una estimación válida del cambio del contrafactual en los resultados del grupo de tratamiento. Este supuesto de que los resultados en los grupos de tratamiento y comparación evolucionan paralelamente a lo largo del tiempo no es siempre posible de probar sin múltiples rondas de datos antes de la intervención. **DD** es particularmente útil en casos en que no se cuenta con datos de panel y el investigador puede establecer un experimento natural que generó una asignación al tratamiento que es aproximadamente aleatoria. En estos casos, existe la gran ventaja de que no se ha requerido implementar el programa de manera aleatoria (con los costos que esto implica) y no es indispensable recolectar datos exclusivamente para la evaluación de impacto o disponer de datos de panel, que son generalmente escasos. Se necesita contar con al menos un par de encuestas de corte transversal que sean representativas de la población elegible para el programa y que contengan los datos indispensables para identificar los grupos de tratamiento y control antes y después de la implementación del programa y demás

variables necesarias para estimar el efecto del programa. La corrección de diferencias en el período pretratamiento permite reducir el sesgo que ocurriría si se observan diferencias preexistentes entre el grupo de tratamiento y el grupo de control. Metodológicamente, es muy sencillo implementar el modelo **DD.** Por supuesto, la desventaja de esta metodología es que no es fácil encontrar eventos que generaron una asignación al programa que es aproximadamente aleatoria Los experimentos naturales son relativamente escasos en la realidad. Si el evento escogido por el investigador como experimento natural en realidad no generó asignación aleatoria (o casi aleatoria) al programa, entonces el modelo de **DD** solo puede eliminar parte del sesgo que se debe a la autoselección para los grupos de tratamiento y control. Por otra parte, se requiere también que en las encuestas disponibles se encuentren los datos indispensables para poder llevar a cabo la evaluación. En algunos casos, el investigador identifica el evento fortuito, pero por ser muy nuevo o muy pequeño el programa, aún no existen los datos en encuestas administrativas que le permitan hacer la evaluación **(Bernal y Peña, 2017).**

- El método **PSM**, tiene los supuestos más fuertes de todos los métodos, y esencialmente descarta cualquier característica no observable entre los participantes del programa y los no participantes. En general, cuanto más fuertes sean los supuestos, mayor será el riesgo de que no se cumplan en la práctica. Este método, es flexible y tiene requerimientos de datos mucho menos exigentes que los demás métodos reseñados en este libro Se puede aplicar a un único levantamiento de información, siempre y cuando exista información para los grupos de tratamiento y control. Por tanto, se usa con mucha frecuencia. Sin embargo, la confiabilidad de sus resultados depende de manera crucial de qué tan factible es que se cumplan los supuestos para el caso de estudio. En particular, los resultados son confiables siempre y cuando existan razones para pensar que las variables no observables o no disponibles en la base de datos, que determinan la variable resultado, no son un determinante de la participación en el programa.

Este supuesto se viola con frecuencia. Este método es preferible al simple modelo de diferencias con regresores adicionales cuando los grupos de tratamiento y control son disímiles, ya que al reducir la comparación a la región de soporte común se disminuye la posibilidad de obtener estimaciones basadas en extrapolaciones que son difíciles de fundamentar. Si bien dista mucho de ser un método ideal/ con buenos datos y decisiones acertadas por parte del investigador, **PSM** puede ser un método poderoso **(Bernal y Peña, 2017).**

- El método de **IV** tiene la gran ventaja de permitir estimaciones del efecto causal de programas con base en datos no experimentales, aun en presencia de variables no observables que determinan tanto la participación en el programa como la variable de resultado; es decir, si se encuentra una variable instrumental que cumpla con las condiciones establecidas en estea obra. Recuerde, por ejemplo, que el método **PSM** discutido requiere el supuesto de que la participación en el programa solo está determinada por variables observables y no por variables no observadas no medidas. Sin embargo, tiene la desventaja de que es difícil encontrar instrumentos validos y, en especial, establecer formalmente la validez de los instrumentos propuestos por el investigador. Por otra parte, es un estimador local, en el sentido de que el impacto se estima para el grupo de individuos cuyas decisiones de participación cambian en respuesta a la **IV** . Esto tiene implicaciones en cuanto al alcance de la interpretación del estimador y las recomendaciones de políticas de los resultados **(Bernal y Peña, 2017).**
- La **RD** depende de la comparabilidad de las unidades justo por encima y justo por debajo del umbral de elegibilidad. Algunas de sus ventajas, incluyen que los supuestos requeridos para implementarlo, sean factibles en la mayoría de los casos, lo cual produce estimadores del impacto del programa que son creíbles; es relativamente fácil de implementar en la práctica y en la actualidad puede ser un diseño frecuente en programas, dado que se ha popularizado el uso de herramientas de focalización de políticas de introducción de inno-

vaciones. Por otra parte, tiene la desventaja de que es un estimador local, en el sentido de que el impacto se estima sólo alrededor del punto de discontinuidad ***Z*,** y esto tiene implicaciones en cuanto al alcance de la interpretación del estimador y las recomendaciones de política de los resultados. Es decir, la metodología es rigurosa en cuanto a su validez interna, pero no respecto a la validez external. Por otra parte, la metodología requiere muestras densas a lado y lado del umbral, lo cual generalmente implica que se requieren muestras muy grandes **(Bernal y Peña, 2017).**

Respaldando la evaluación

Es útil tener un plan de respaldo en caso de que la primera opción de metodología no funcione. Por ejemplo, aún realizando una planeación estratégica adecuada, siempre se epresentan situaciones repentinas, (como la pandemia mundial COVID-19, de principios de 2020), en las que un programa de innovación , no cumple con las solicitudes previstas dada la ralentizaciób de la economía. También esto suele ocurrir con cambios no anticipados en el contexto operativo o político de un programa. La planificación es una buena práctica desde un punto de vista metodológico, en el uso de las técnicas de evaluación de impacto. Al tener dudas acerca de si algún método muestra sesgos, se podrán verificar los resultados comparándolos con el otro método.

Buscando la unidad de intervención mínima

Las reglas operativas, frecuentemente, determinan el nivel en que se puede asignar una intervención, algo que se relaciona con la manera en que se implementa el programa. Por ejemplo, si se pone en marcha un programa de innovación orientada a mejor la educación a nivel de distrito, todas las comunidades del distrito o recibirían el programa (como grupo) o no lo

recibirán. equipos de captación. Existen casos donde los programas se implementan de manera más eficiente, a nivel individual o pyme, mientras que otros deben aplicarse a nivel de la sector o a un nivel administrativo superior. Incluso si un programa se asigna e implementar a nivel individual, el equipo de administración de la innovación, quizá requiera de un nivel superior de agregación a fin de mitigar los *efectos potenciales de derrame* (*spillover*); es decir, los *efectos indirectos* de las unidades que participan en las unidades que no participan. Implementar una intervención a un mayor nivel tiende a generar problemas para la evaluación, por dos motivos:

- En primer lugar, las evaluaciones de las intervenciones asignadas e implementadas a niveles superiores, como el sector o el distrito administrativo, *requieren tamaños de muestra más grandes* y por tantos, serán más costosas, en comparación con las evaluaciones de intervenciones a un nivel más bajo, como el nivel individual o de pyme. El nivel de intervención es importante porque define la unidad de asignación a los grupos de tratamiento y comparación, y eso tiene implicaciones para el tamaño de la muestra de la evaluación y su costo. En implementación de intervenciones a nivel superior, se necesita una muestra más grande que tiende a crear *conglomerados* (*clusters*) para detectar el impacto del programa.
- En segundo lugar, a niveles superiores de intervención, es más difícil encontrar un número suficiente de unidades para realizar la evaluación. Sin embargo, la asignación aleatoria solo genera grupos de tratamiento y comparación comparables si se lleva a cabo con un número suficiente de unidades. Por ejemplo, si el nivel de agregación es por región y el estado solo tiene seis regiones, existe baja probilidad que la *aleatorización* genere equilibrio entre los grupos de *tratamiento y comparación*. Esto es porque la clave para equilibrar ambos grupos, es el número de unidades asignadas, no el número de individuos o de pymes de la muestra. Por lo tanto, llevar a cabo una asignación aleatoria en niveles altos de implementación pone en riesgo la *validez interna* si el número de unidades no es suficiente.

Cómo lograrlo

Con la implementación de la intervención en un nivel geográfico administrativo alto, el equipo de evaluación y los administradores del programa de introducción de innovaciones deben trabajar en encontrar la unidad de intervención más pequeña que sea operacionalmente factible; con esto, se evitan los riesgos asociados. Diversos factores, se deben considerar para determinar la unidad de intervención más pequeña factible:

- Las economías de escala y la complejidad administrativa en la implementación del programa.
- La capacidad administrativa para asignar beneficios a nivel individual o de los hogares.
- Preocupaciones potenciales a propósito de posibles tensiones.
- Preocupaciones potenciales acerca de los efectos de derrame y la contaminación del grupo de comparación.

Las economías de escala hacen que dependa la unidad factible de intervención más pequeña así como la complejidad administrativa de realizar el programa. Por ejemplo, un programa de innovación agrícola, quizá requiera una oficina local para que los beneficiarios presenten reclamos y para pagar a los proveedores. Los costos fijos de la oficina tienen que repartirse entre un gran número de beneficiarios, de modo que puede ser ineficiente implementar el programa a nivel individual y más eficiente si se hace a nivel de la región.

Sin embargo, en situaciones con tipos de intervenciones nuevas y no probadas, se deberá considerar absorber las ineficiencias de corto plazo e implementar el programa en los distritos administrativos, para asegurar la credibilidad de la evaluación y disminuir los costos de la recopilación de datos.

Algunos administradores de programas de innovación, afirman que los programas administrados a nivel local, como los programas de innovación de salud o educación, no tienen las capacidades administrativas para implementar programas a nivel individual. Su preocupación es que

sería una carga crear sistemas para prestar diferentes beneficios a diferentes beneficiarios en unidades administrativas locales, y que acaso resulte difícil garantizar que la asignación a los grupos de tratamiento y comparación se implemente siguiendo el diseño.

Para una evaluación de impacto, esto último representa una seria amenaza dado que los administradores del programa de innovación, quizá no puedan poner en marcha el programa de forma consistente siguiendo un diseño de evaluación. En este caso, se hace necesario o una implementación a un nivel superior o una simplificación del diseño de evaluación de impacto.

Los gobiernos en ocasiones, prefieren implementar programas a niveles más agregados, como el de la comunidad o región ya que les preocupan las tensiones potenciales que surgen cuando los miembros de los grupos de comparación observan que los vecinos en el grupo de tratamiento tienen derecho a los beneficios. Numerosos programas se han llevado a cabo con éxito a nivel individual o pyme en las comunidades sin generar tensiones, sobre todo cuando los beneficios se han asignado de manera equitativa, transparente y sujetos a rendición de cuentas. Sin embargo, se recomienda siempre tener en cuenta el riesgo de la posibilidad de que surjan tensiones en el contexto de una evaluación de impacto específica.

Por último, en el surgimiento e implementación de un programa de innovación a niveles muy bajos, como sujetos, o pymes, la contaminación del *grupo de comparación* tiende a poner en entredicho *la validez interna* de la evaluación. Un caso de ejemplo es suponer, que se evalúa el efecto de proporcionar innovaciones que faciliten la eliminación de aceites de desperdicio para la mejora de la salud de los hogares. Si se instalan equipos de captación para un hogar pero no para su vecino, el hogar de tratamiento bien puede compartir el uso de los equipos de captación, con un *vecino de comparación* y, por lo tanto, el hogar vecino no sería una verdadera comparación, dado que se beneficiaría del *efecto de derrame (spillover).*

En la práctica, los administradores de programas de introducción de innovaciones, tienen que optar por la unidad factible de intervención más pequeña que:

- Permita contar con un gran número de unidades en la evaluación.
- Mitigue los riesgos para la validez interna, y
- Se ajuste al contexto operativo.

La metodología elegida y cómo evitar sus dificultades de implementación

Como se podrá apreciar, la mayoría de los métodos de evaluación de impacto producen estimaciones válidas del contrafactual si se encuentran con supuestos muy concretos o específicos. Cualquier método, asume varios riesgos:

- El que los supuestos fundamentales que adopte, no sean válidos, lo que produce resultados de estimación sesgados, del impacto del programa que se evalúa.
- Un tipo de riesgo surge cuando se estima el impacto de un programa en todo un grupo y los resultados ocultan algunas diferencias en las respuestas al tratamiento de los diferentes receptores, es decir, los *efectos heterogéneos del tratamiento*.
- El supuesto de que la mayoría de los métodos de evaluación de impacto que un programa, influye en los resultados de una manera simple y lineal para todas las unidades de la población. Sin embargo, se debe considerar que diferentes subpoblaciones, pueden tener un impacto del programa de manera muy diferente, que convenga tener muestras estratificadas para cada subpoblación. Por ejemplo, suponga que al investigador le interesa conocer el impacto de **pei** de **pbt** de diseño de software, pero solo el **10%** de las **pbt** de la región está conformado por éstas. En ese caso, es posible que incluso, una

muestra aleatoria *grande* de **pbt** no contenga un número suficiente de las de diseño de software y estimar el impacto del programa en las **pbt**. Para el diseño de la muestra de evaluación, convendrá estratificar la misma basándose en el tipo de actividad, e incluir un número suficientemente grande de **pbt** de diseño de software a fin de detectar un determinado tamaño del efecto.

La conducta y sus efectos no intencionados

Gertler et al. (2017) afirman que, al llevar a cabo una evaluación de impacto, es neceario reconocer la inducción a *respuestas no intencionadas en la conducta* de la población que se estudia, los cuales citamos en la **Tabla 8.4.**

Tabla 8.4. Efectos no intencionados de la conducta

Efecto	Descripción
Hawthorne	Ocurre cuando debido al mero hecho de saber que están siendo observadas, las unidades se comportan de manera diferente.
John Henry	Se produce cuando las unidades de comparación se esfuerzan más para compensar el hecho de no ser objeto del tratamiento.
Anticipación	En una aleatorización por fases, es posible que las unidades del grupo de comparación, esperen recibir el programa en el futuro y comiencen a cambiar su comportamiento antes de que el programa realmente se materialice.
Sesgo por sustitución	Influye en el grupo de comparación: las unidades que no fueron seleccionadas para ser objeto del programa pueden encontrar buenos sustitutos gracias a su propia iniciativa.

Fuente: Gertler et al. (2017) con adaptación propia.

Un factor que tiende a socavar la validez interna de los resultados de la evaluación, aunque se use la asignación aleatoria como método de evaluación, son las respuestas en la conducta que afectan de manera despropor-

cionada al grupo de comparación. Esto constituye un problema porque un grupo de comparación que se esfuerza más para compensar el hecho de no ser objeto de un tratamiento, o que cambia su conducta en previsión del programa, no es una buena representación del contrafactual.

Si se tiene algún motivo para creer que se pueden producir estas respuestas no intencionadas en la conducta, a veces una opción es constituir otros grupos de comparación que no se vean en absoluto afectados por la intervención, es decir, un grupo que permita explícitamente comprobar dichas respuestas. También es posible, adoptar como buena práctica, el recopilar datos cualitativos con el fin de entender mejor las respuestas en la conducta.

Imperfección del cumplimiento

La *imperfección del cumplimiento* es la diferencia entre la condición asignada del tratamiento y la condición real del mismo. La *imperfección del cumplimiento* es generada en el momento en que, *cuando algunas unidades asignadas al grupo de tratamiento no reciben tratamiento*, y cuando *algunas unidades asignadas al grupo de comparación reciben tratamiento*. En Variables Instrumentales (**IV**. Capítulo 6) se estudia la imperfección del cumplimiento en referencia a la asignación aleatoria, si bien la imperfección del cumplimiento también se puede producir con el diseño de regresión discontinua (**RD**. Capítulo 8) y con doble diferencia (**DD**. Capítulo 5). Antes de interpretar las estimaciones de impacto que genera cualquier método, se debe saber si se ha producido una imperfección del cumplimiento en el programa. **Ver Tabla 8.5.**

Tabla 8.5. Imperfección del cumplimiento

Descripción
No todos los participantes previstos participan realmente en el programa. A veces, algunas unidades asignadas a un programa deciden no participar.
Algunos participantes previstos son excluidos del programa debido a errores administrativos o de ejecución.
Se ofrece el programa por error a algunas unidades del grupo de comparación, que se inscriben en él.
Algunas unidades del grupo de comparación consiguen participar en el programa a pesar de que no se les ofrece.
El programa se asigna a partir del índice continuo de elegibilidad, pero no se aplica estrictamente el umbral de elegibilidad.
Se produce una *migración selectiva* en función de la condición del tratamiento. Por ejemplo, puede que la evaluación compare los resultados en las unidades tratados y no tratados, pero las personas pueden decidir trasladarse a otra unidad si no les agrada la condición de tratamiento de esa unidad.

Fuente: Gertler et al. (2017) con adaptación propia.

Frecuentemente ante la imperfección del cumplimiento, los métodos de evaluación de impacto, producen estimaciones de la intención de tratar. Sin embargo, mediante el método de variables instrumentales, es posible obtener *estimaciones del tratamiento en los tratados* a partir de las *estimaciones de la intención de tratar.* En el capítulo 6 (**IV**) se explicó la idea básica para abordar *imperfección del cumplimiento en el contexto de la asignación aleatoria*. A través de ajustar en la muestra de la evaluación, el porcentaje de cumplidores, es posible recuperar el efecto local promedio del tratamiento en los cumplidores a partir de la *estimación de la intención de tratar*. Mediante la aplicación del enfoque más general de variables instrumentales, dicho *ajuste* es posible extenderlo a otros métodos. La variable instrumental contiene una fuente externa de variación que ayuda a eliminar o corregir el sesgo que se genera de la imperfección en el cumplimiento. En el caso de la asignación aleatoria con imperfección en el cumplimiento, se utiliza una variable 0/1 (*dummy*) que asume el valor de 1 si la unidad

estaba asignada originalmente al grupo de tratamiento, y de 0 si la unidad estaba originalmente asignada al grupo de comparación. Durante la fase de análisis, la **IV** se usa con frecuencia en el contexto de una *regresión en dos fases* que identifica el impacto del tratamiento en los cumplidores. **Ver Tabla 8.6.**

Tabla 8.6. Imperfección del cumplimiento usando variable instrumental (IV)

Descripción
En el contexto del diseño de regresión discontinua, debe utilizarse una variable 0/1 que indique si la unidad se encuentra en el lado no elegible o elegible de la puntuación límite.
En el contexto de la migración selectiva, una posible variable instrumental para la ubicación del individuo después del comienzo del programa sería la ubicación del individuo antes del anuncio del programa.
Desde un punto de vista técnico, no es deseable que una gran parte del grupo de comparación se inscriba en el programa. A medida que aumenta la proporción del grupo de comparación que se inscribe en el programa, la fracción de *cumplidores* en la población disminuirá, y el efecto local promedio del tratamiento estimado con el método de variable instrumental será válido solo para una fracción cada vez más pequeña de la población de inter.é. Si esto se extiende demasiado, es posible que los resultados pierdan toda relevancia para las políticas de introducción de innovaciones, dado que ya no serían aplicables a una parte suficientemente grande de la población de interés.
Tampoco es deseable que una parte grande del grupo de tratamiento siga sin inscribirse. Una vez más, a medida que la fracción del grupo de tratamiento que se inscribe en el programa disminuye, también lo hace la fracción de *cumplidores* de la población. El efecto promedio del tratamiento estimado con el método de variable instrumental será válido solo para una fracción cada vez menor de la población de interés.
Como ya se trató en el capítulo 6, el método de variables instrumentales es válido solo en ciertas circunstancias; decididamente no es una solución universal.

Fuente: Gertler et al. (2017) con adaptación propia.

Efectos indirectos o de derrame (*spillover*)

Los efectos indirectos o de derrames son otro problema habitual a los que se enfrentan las evaluaciones de impacto, sea que se aplique el método de asignación aleatoria, el de doble diferencia o el de regresión discontinua. Se produce cuando una intervención afecta a un no participante, y puede ser positivo o negativo. Hay cuatro tipos de efectos de derrame **(Angelucci y Di Maro 2015). Ver Tabla 8.7.**

Tabla 8.7. Tipos de efectos indirectos o de derrame (*spillovers*)

Externalidades
Se trata de efectos que van de los *sujetos tratados a los sujetos no tratados*. Por ejemplo, en un proceso de innovación de innoculación de vacunas en un pueblo, reduce la probabilidad de que los habitantes no vacunados del mismo pueblo, contraigan esa enfermedad. Se trata de un ejemplo de externalidades positivas. Las externalidades también pueden ser negativas. Por ejemplo, los cultivos de un proceso de innovación que aplique un agricultor, pueden verse parcialmente destruidos si su vecino aplica un herbicida en su propio terreno y parte del herbicida cae sobre el otro lado de la línea divisoria de la propiedad.
Interacción social
Los efectos de derrame pueden ser el producto de interacciones sociales y económicas entre *poblaciones tratadas y no tratadas*, que conducen a impactos indirectos en los no tratados. Por ejemplo, un estudiante que recibe una Tablet como parte de un programa de innovación de aprendizaje, puede compartir el dispositivo con otro alumno que no participa en el programa.
Efectos de equilibrio del contexto
Estos efectos se producen cuando una intervención influye en las normas comportamentales o sociales dentro de un determinado contexto, como una localidad tratada. Por ejemplo, aumentar la cantidad de recursos de innovación que reciben los **pbt** orientados a la salud tratados de manera que puedan ampliar su gama de servicios, puede influir en las expectativas de la población a propósito de cuál debería ser el nivel de los servicios ofrecidos en todos **pbt** orientados a la salud.

Efectos de equilibrio general
Estos efectos se producen cuando las intervenciones influyen en la oferta y demanda de bienes y servicios y, por ende, cambian el precio de mercado de esos servicios. Por ejemplo, un programa que entrega vales a las **pbt** con CEO femeninas pobres para que utilicen los centros públicos de innovación para realizar pruebas de sus prototipos, tiene la posibilidad de aumentar la demanda de servicios en los centros públicos, lo que incrementaría el precio del servicio para todos.

Fuente: Angelucci y Di Maro (2015) con adaptación propia.

Si el no participante que experimenta el derrame, pertenece al *grupo de comparación*, el efecto derrame viola el requisito básico de que el resultado de una unidad no debería verse afectado por la asignación concreta de tratamientos a otras comunidades. Este supuesto de estabilidad del valor de la unidad de tratamiento (**SUTVA**. *Stable Unit Treatment Value Assumption*) es necesario para asegurar que la asignación aleatoria produzca estimaciones no sesgadas del impacto. Si el grupo de control se ve indirectamente afectado por el tratamiento recibido por el grupo de tratamiento (por ejemplo, estudiantes del *grupo de comparación* que usan las Tablets de los estudiantes del *grupo de tratamiento*), la comparación no representa con precisión qué habría ocurrido en el *grupo de tratamiento* en ausencia de *tratamiento* (el contrafactual).

Si el no participante que experimenta el derrame no pertenece al grupo de comparación, el supuesto **SUTVA** sería válido y el grupo de comparación seguiría proporcionando una buena estimación del contrafactual. Sin embargo, aún habría que medir el derrame, porque representa un impacto real del programa. En otras palabras, la comparación de los resultados de los *grupos de tratamiento y comparación* generaría *estimaciones no sesgadas* del impacto del tratamiento en el grupo tratado, pero esto no tendría en cuenta el impacto del programa en otros grupos.

Sugerencias de diseño

Si se encuentra en el proceso de diseño de una evaluación de impacto para un programa donde es probable que se produzcan efectos indirectos o derrames, debe priorizar que el objetivo de la evaluación necesita ser más amplio. Mientras que una evaluación de impacto estándar, su objetivo es estimar el impacto (o efecto causal) de un programa de innovación, en un resultado de interés para las unidades que reciben el tratamiento, una evaluación con potenciales efectos externos de derrame, tendrá que responder a dos preguntas (**ver Tabla 8.8**).

Tabla 8.8. Preguntas a realizar ante diseño de evaluación de impacto con efectos indirectos o de derrame (*spillovers*)

Descripción
Modalidad de Impacto Directo: ¿Cuál es el impacto (o efecto causal) de un programa en un resultado de interés para las unidades que reciben el tratamiento? Se trata del impacto directo que el programa tiene en los grupos tratados. Para estimar el impacto directo en los grupos tratados, habrá que elegir el grupo de comparación de tal manera que no se vea afectado por los derrames. Por ejemplo, puede ponerse como condición que las **pbt**, pueblos, clínicas u hogares de tratamiento y comparación están situados lo suficientemente lejos unos de otros de manera que los derrames sean poco probables.
Modalidad de Impacto Indirecto: ¿Cuál es el impacto (o efecto causal) de un programa en un resultado de interés en las unidades que no reciben el tratamiento? Se trata del impacto indirecto. que el programa tiene en los grupos no tratados. Para estimar el impacto indirecto en los grupos no tratados, debería identificarse para cada grupo no tratado un grupo de comparación adicional que pueda verse afectado por los derrames. Por ejemplo, los trabajadores comunitarios de u proceso de innovación de la salud pueden realizar visitas domiciliarias para proporcionar información a los padres acerca de los beneficios de una dieta variada mejorada para los niños. Supóngase que dichps trabajadores, solo visitan algunos hogares de un pueblo determinado. Uno puede estar interesado en los efectos de derrame sobre los niños de los hogares no visitados, en cuyo caso necesitaría hallar un grupo de comparación para estos niños. Al mismo tiempo, puede ser que la intervención también afecte la variedad de la dieta de los adultos. Si tal efecto indirecto es de interés para la evaluación, se necesitaría también un grupo de comparación para los adultos. A medida que aumente el número de canales potenciales de derrame, el diseño puede complicarse con relativa rapidez.

Fuente: Gertler et al. (2017) con adaptación propia.

Las evaluaciones con efectos indirectos o de derrame (*spillovers*), plantean ciertos problemas específicos:

- En primer lugar, cuando los efectos de derrame son *probables*, *es importante entender el mecanismo de derrame*, ya sea biológico, social, ambiental o de otro tipo. Si no se sabe cuál es el mecanismo de derrame, no será posible elegir con precisión los grupos de comparación que son y no son afectados por los derrames.
- En segundo lugar, una evaluación con efectos de derrame requiere una recopilación de datos más amplia que una evaluación en la cual esa preocupación no existe: *hay un grupo de comparación adicional* (en el ejemplo anterior, los pueblos vecinos). Puede que también tengan que recopilarse datos sobre las otras unidades (en el ejemplo anterior, los adultos de los hogares objetivo para visitas relacionadas con la nutrición de los niños).

El desgaste

Otro problema habitual que afecta a las evaluaciones de impacto, es el *sesgo del desgaste* ya sea con el método de asignación aleatoria, de regresión discontinua (**RD**) o de doble diferencia (**DD**). El *desgaste* ocurre cuando partes de la muestra, *desaparecen* a lo largo del tiempo y los investigadores no pueden encontrar a todos los miembros iniciales de los grupos de tratamiento y comparación en las encuestas o en los datos de seguimiento. Por ejemplo, de los **2000 pbt** encuestados en la línea de base, los investigadores pueden encontrar solo **1800** en una encuesta de seguimiento dos años después. Si intentan volver a realizar la encuesta al mismo grupo, por ejemplo, **10** años después, es posible que encuentren incluso menos **pbt** originales. El *desgaste* se puede producir por diferentes motivos:

- Por ejemplo, puede que los miembros de los pbt, hogares o incluso familias enteras se muden a otro pueblo, ciudad, región, o incluso país.
- En otros casos, los encuestados ya no están dispuestos a responder a una segunda encuesta.
- También ocurre que los conflictos y la falta de seguridad en la zona impidan que el equipo de investigación lleve a cabo una encuesta en algunas localidades incluidas en la línea de base.

Lo que ocasiona como problema

El desgaste es problemático por dos motivos:

- En primer lugar, la muestra de seguimiento quizá ya no represente adecuadamente a la población de interés. Recuerde que cuando se elige la muestra, en el momento de la *asignación aleatoria*, se hace de manera que represente de forma apropiada a la población de interés. Es decir, se escoge una muestra que tiene validez externa para la población de interés. Si la encuesta o la recopilación de datos de seguimiento se ve limitada por un desgaste considerable, debería ser preocupante que la muestra de seguimiento represente solo a un subconjunto específico de la población de interés. Por ejemplo, si las personas de mayor nivel educativo de la muestra original también son las que emigran, la encuesta de seguimiento ignoraría a aquellas personas con estudios y ya no representaría adecuadamente a la población de interés, que incluía a esas personas.
- En segundo lugar, es posible que la muestra de seguimiento ya no esté equilibrada entre el grupo de tratamiento y de comparación. Supóngase que se intenta evaluar un programa de innovación para mejorar la educación de las niñas y que es más probable que las niñas con estudios se muden a la ciudad a buscar un empleo. Entonces, la encuesta de seguimiento podría mostrar un alto desgaste desproporcionado en el *grupo de tratamiento*, en relación con el *grupo de*

comparación. Esto podría afectar la validez interna del programa, es decir, al contrastar las unidades de tratamiento y comparación que se encuentran en el seguimiento, ya no se podrá dar una estimación precisa del impacto del programa.

Cómo superarlo

Si durante las encuestas de seguimiento se halla desgaste, los siguientes dos pasos pueden ayudar a evaluar el alcance del problema **(Gertler et al., 2017):**

- Primero, verifique si las características de *línea de base* de las unidades que abandonaron la muestra son estadísticamente iguales a las características de línea de base de las unidades que fueron encuestadas con éxito la segunda vez. Siempre que las características de *línea de base* de ambos grupos no sean estadísticamente diferentes, la nueva muestra debería seguir representando a la población de interés.
- Segundo, verifique si la *tasa de desgaste* del grupo de tratamiento es similar a la *tasa de desgaste* del grupo de comparación. Si ambas son significativamente diferentes, surge la preocupación de que la muestra ya no sea válida, y quizá deban utilizarse diversas técnicas estadísticas para intentar corregir esto. Un método habitual es la *ponderación por probabilidad inversa*, un método que repondera estadísticamente los datos (en este caso, los datos de seguimiento) para corregir el hecho de que una parte de los encuestados originales está ausente. El método formula una reponderación de la muestra de seguimiento de modo que tenga un aspecto similar a la muestra de *línea de base* **(Lee, 2009).**

Los efectos y su persistencia en el tiempo

Los canales de transmisión entre insumos, actividades, productos y resultados suelen estar estrechamente relacionados con los cambios en el comportamiento humano ya que tienden o a ocurrir de inmediato, pronto o después de un período de tiempo. En el capítulo 2 se expone la importancia de pensar en estos canales para planificar antes de que comience la intervención y desarrollar una cadena causal clara para el programa que se esté evaluando. Es fundamental considerar estos aspectos cuando se diseña una evaluación respecto del tiempo, porque:

- Los programas no se vuelven plenamente efectivos justo después de su inicio **(King y Behrman, 2009).** Los administradores de un programa de introducción de innovaciones, necesitan tiempo para que éste comience a funcionar, y es posible que los beneficiarios no vean los frutos de inmediato porque los cambios de conducta requieren tiempo, y las instituciones tampoco modifican su comportamiento con rapidez.
- Una vez que las instituciones y los beneficiarios cambian ciertas conductas, existe la posibilidad de que estas se mantengan aun cuando se suspenda el programa. Por ejemplo, un programa que incentiva a las **pbt** a separar y reciclar la basura y ahorrar energía puede seguir siendo efectivo después de que se eliminen los incentivos, si consigue cambiar las normas de las **pbt** en el manejo de la basura y la energía.
- Cuando se diseña una evaluación, hay que tener mucho cuidado (y ser realistas) para definir cuánto podría tardar el programa en alcanzar su plena efectividad. Es posible, que sea necesario llevar a cabo diversas encuestas de seguimiento para medir el impacto del programa a lo largo del tiempo, o incluso después de que el programa se interrumpa.

Combinando diversas opciones de tratamiento en la evaluación de impacto

La realidad suele ser más compleja que tener análisis de programas de innovación con un solo tipo de tratamiento. Sin embargo, muchas cuestiones relevantes relacionadas con las políticas se plantean en programas multifacéticos, es decir, que combinan varias opciones de tratamiento **(Banerjee y Duflo, 2009)**. Al diseñar las políticas de introducción de innovaciones, se suele requerir saber si funciona mejor o tiene un costo menor que otro programa, no solo si el programa funciona o no. Por ejemplo, si se quiere aumentar la introducción de innovaciones tecnológicas que fomenten las clases virtuales con la adopción de nuevas tecnologías de información y comunicaciones, ¿es más eficaz orientar las intervenciones a la demanda (como las transferencias condicionadas a las familias) o a la oferta (como mayores incentivos para los profesores)? Y si se introducen las dos intervenciones conjuntamente, ¿funcionan mejor que cada una por su cuenta?, ¿son complementarias? Si la relación costo-efectividad es una prioridad, cabe preguntarse también, ¿cuál es el nivel óptimo de los servicios que debe prestar el programa. Por ejemplo, ¿cuál es la duración óptima de un programa de capacitación para el empleo? ¿programas de capacitación de **4** meses contribuye más que un programa de dos meses? De ser así, ¿la diferencia es lo suficientemente grande para justificar los recursos adicionales necesarios para un programa de cuatro meses? Por último, los administradores de introducción de innovaciones tienden a interesarse en cómo alterar un programa existente para hacerlo más efectivo, y probar diversos mecanismos con el fin de encontrar cuál(es) funciona(n) mejor.

Las evaluaciones de impacto pueden ayudar a responder preguntas más generales, además de estimar el impacto de una intervención sobre un resultado de interés, como las siguientes (**ver Tabla 8.9**).

Tabla 8.9. Preguntas a realizar ante diseños de intervención de más de una técnica de impacto

Descripción
¿Cuál es el impacto de un tratamiento en comparación con otro? Por ejemplo, ¿cuál es el impacto en el desarrollo cognitivo de los niños de un programa de innovación de procesos que ofrece capacitación a los padres, en comparación con una intervención de innovación sobre el procesos de nutrición?
¿El impacto conjunto de un primer y un segundo tratamiento es mayor que la suma de los dos impactos? Por ejemplo, ¿el impacto de la intervención en un proceso de innovación de aprendizaje basado en tecnologías de información con capacitación de padres y la intervención sobre procesos de innovación enfocados a la nutrición es mayor, menor o igual que la suma de los efectos de cada una de las intervenciones?
¿Cuál es el impacto de un tratamiento de alta intensidad en comparación con un tratamiento de menor intensidad? Por ejemplo, ¿cuál es el efecto en el desarrollo de la intervención en un proceso de innovación de aprendizaje basado en tecnologías de información cognitivo de niños con retraso en el crecimiento si un trabajador social los visita en su casa cada dos semanas, en lugar de visitarlos una vez al mes?

Fuente: Gertler et al. (2017) con adaptación propia.

En el caso de diseños de evaluaciones de impacto para dos tipos de programas multifacéticos, es posible que existan:

- Los que tienen múltiples niveles del mismo tratamiento, y
- Los que tienen múltiples tratamientos.

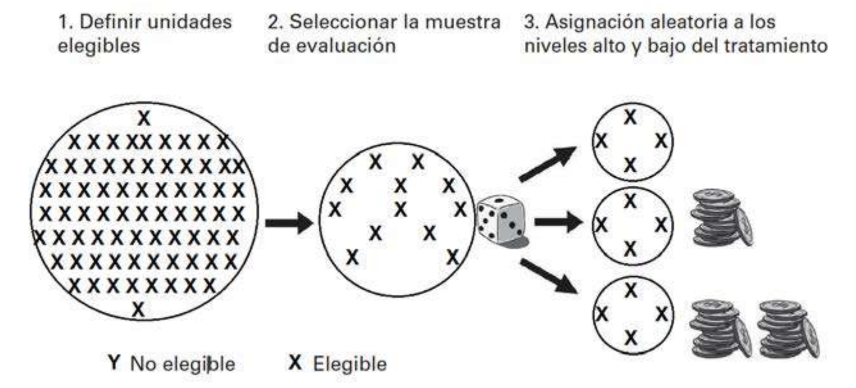
En estos casos:

- Primero, se analiza cómo diseñar una evaluación de impacto de un programa con varios niveles de tratamiento.
- Después, se examinan los diferentes tipos de impactos de un programa con múltiples tratamientos. Para este análisis se supone que debe usarse un método de asignación aleatoria, aunque puede generalizarse a otros métodos.

Diferentes niveles de tratamiento

Es relativamente fácil, diseñar una evaluación de impacto para un programa con niveles variables de tratamiento. Por ejemplo, suponga que se intenta evaluar el impacto de un programa con dos niveles de tratamiento: alto (por ejemplo, visitas cada dos semanas) y bajo (visitas mensuales). Se requiere evaluar el impacto de ambas opciones, y saber cuánto afectan a los resultados esas visitas adicionales. Para ello, se puede organizar un *sorteo* de modo de decidir quién recibe el nivel alto de tratamiento, quién recibe el nivel bajo de tratamiento y a quién se asigna al grupo de comparación. **Ver Figura 8.3.**

Figura 8.3. Pasos para la asignación aleatoria de dos niveles de tratamiento



Fuente: Gertler et al. (2017) con adaptación propia.

Como se sabe, en la *asignación aleatoria*:

- El primer paso consiste en definir la población de unidades elegibles para el programa.
- El segundo, en seleccionar una muestra aleatoria de unidades que se incluirá en la evaluación, la denominada muestra de evaluación.
- Una vez que se cuente con la muestra de evaluación, en el tercer paso se asignarán aleatoriamente unidades al grupo que recibe un nivel

alto de tratamiento, al grupo que recibe el nivel bajo de tratamiento o al grupo de comparación.

- Como resultado de la asignación aleatoria a múltiples niveles de tratamiento, se habrán creado tres grupos distintos:
  - El **grupo A** es el grupo de comparación.
  - El **grupo B** recibe el nivel bajo de tratamiento.
  - El **grupo C** recibe el nivel alto de tratamiento.

Cuando se implementa correctamente, la *asignación aleatoria* garantiza que los *tres grupos sean similares*. Por lo tanto, se puede estimar el impacto del nivel alto de tratamiento mediante la comparación del resultado promedio del grupo **C** con el resultado promedio del grupo **A**. También se puede estimar el nivel bajo de tratamiento comparando el resultado promedio del grupo **B** con el del grupo **A**. Finalmente, se tiene la posibilidad de evaluar si el nivel alto de tratamiento tiene un mayor impacto que el nivel bajo de tratamiento comparando los resultados promedio de los grupos **B** y **C**. La estimación del impacto de un programa con más de dos niveles de tratamiento seguirá la misma lógica. Si existen tres niveles de tratamiento, el proceso de asignación aleatoria creará tres grupos de tratamiento diferentes, además de un grupo de comparación. En general, con ***n*** niveles de tratamiento, habrá ***n*** *grupos de tratamiento*, más un *grupo de comparación.*

Evaluación de intervenciones múltiples

Además de comparar varios niveles de tratamiento, también se pueden comparar *opciones de tratamiento totalmente diferentes*. De hecho, los responsables de las políticas de introducción de innovaciones, prefieren comparar los méritos relativos de diferentes intervenciones, más que conocer solo el impacto de una intervención. Suponga que se propone evaluar el impacto en la adopción de un nuevo proceso para fabricación de agua de coco de un programa, con dos intervenciones, transferencias condi-

cionadas a los fabricantes y distribución gratuita en autobús a los centros de distribución. Así que:

- Primero, debe conocer el impacto de cada intervención por separado. Este caso es prácticamente idéntico a aquel en que se prueban diferentes niveles de tratamiento de una intervención, a saber, en lugar de asignar aleatoriamente las unidades a niveles altos y bajos de tratamiento y al grupo de comparación, se les puede asignar de forma aleatoria a un grupo de transferencias condicionadas, a un grupo de transporte gratuito en autobús y al grupo de comparación. En general, con ***n*** niveles de tratamiento, habrá ***n*** grupos de tratamiento, más un grupo de comparación. Aparte de querer conocer el impacto de cada intervención por separado, puede que también se desee conocer si la combinación de los dos es mejor que la simple suma de los efectos individuales. Desde el punto de vista de los participantes, el programa está disponible en tres formas diferentes: solo transferencias condicionadas, únicamente transporte gratuito en autobús o una combinación de transferencias y transporte gratuito. La asignación aleatoria para un programa con dos intervenciones es muy similar al proceso de un programa con una sola intervención. La principal diferencia es la necesidad de organizar varios sorteos independientes, en lugar de uno. Esto produce un diseño cruzado, a veces llamado *diseño transversal*. En la **Figura 8.4** se ilustra este proceso.

Figura 8.4. Pasos para la asignación aleatoria de dos intervenciones

1. Definir unidades elegibles
2. Seleccionar la muestra de evaluación
3. Asignación aleatoria al primer tratamiento
4. y 5. Asignación aleatoria al segundo tratamiento

Y No elegible
X Elegible

Fuente: Gertler et al. (2017) con adaptación propia.

Como en el caso anterior:

- En el primer paso se define la población de unidades elegibles para el programa.
- Segundo paso consiste en seleccionar una muestra aleatoria de unidades elegibles para formar la muestra de evaluación. Una vez obtenida la muestra de evaluación, se procede al
- Tercer paso se asignan aleatoriamente sus unidades a un grupo de tratamiento y a un grupo de control.
- En el cuarto paso, se lleva a cabo un segundo sorteo para asignar de forma aleatoria una subserie del grupo de tratamiento a fin de que reciba la segunda intervención.
- Por último, en el quinto paso se realiza otro sorteo para asignar una subserie del grupo de comparación inicial a fin de que reciba la segunda intervención, mientras que la otra subserie se mantiene como un conjunto puro de comparación.

Como consecuencia de la asignación aleatoria a los dos tratamientos, se habrán creado cuatro grupos, como se muestra en la **Figura 8.5** se muestra este proceso:

Figura 8.5. Diseño híbrido para un programa con dos intervenciones

Fuente: Gertler et al. (2017) con adaptación propia.

- El grupo **A** recibe ambas intervenciones (transferencias condicionadas y transporte en autobús).
- El grupo **B** recibe la primera intervención pero no la segunda (solo transferencia condicionadas).
- El grupo **C** no recibe la primera intervención pero sí la segunda (solo el transporte en autobús)
- El grupo **D** no recibe ni la primera ni la segunda intervención, y constituye el grupo de comparación puro.

Cuando se implementa correctamente, la asignación aleatoria garantiza que los cuatro grupos sean similares. Por lo tanto, se puede estimar el impacto de la primera intervención comparando el resultado del grupo **B** (por ejemplo, la tasa de asistencia escolar) con el resultado del grupo puro

de comparación, el grupo **D**. También se puede estimar el impacto de la segunda intervención comparando el resultado del grupo **C** con el resultado del grupo de comparación puro, el grupo **D**. Además, este diseño también permite comparar el impacto progresivo de recibir la segunda intervención cuando una unidad ya ha recibido la primera. La comparación de los resultados del grupo **A** y del grupo **B** determinará el impacto de la segunda intervención para aquellas unidades que ya han recibido la primera intervención. La comparación de los resultados de los grupos **A** y **C** determinará el impacto de la primera intervención en las unidades que ya han recibido la segunda intervención.

En la descripción anterior se ha usado el ejemplo de la asignación aleatoria para explicar la manera de diseñar una evaluación de impacto para un programa con dos intervenciones diferentes. Cuando un programa cuenta con más de dos intervenciones, se puede aumentar el número de sorteos y continuar subdividiendo la evaluación para formar grupos que reciben las diversas combinaciones de intervenciones. También se pueden implementar múltiples tratamientos y múltiples niveles de tratamiento. Aunque se amplíe el número de grupos, la teoría fundamental del diseño sigue siendo la misma que la descripta anteriormente.

Sin embargo, la evaluación de más de una o dos intervenciones generará dificultades prácticas tanto en la evaluación como en el funcionamiento del programa, ya que la complejidad del diseño incrementará exponencialmente el número de ramas de tratamiento:

- Para evaluar el impacto de una sola intervención se necesitan únicamente dos grupos, uno de tratamiento y otro de comparación.
- Para evaluar el impacto de dos intervenciones se necesitan cuatro grupos, tres de tratamiento y uno de comparación.
- Si se quisiera evaluar el impacto de tres intervenciones, incluidas todas las combinaciones posibles entre ellas, se necesitaría **2 × 2 × 2 = 8** grupos en la evaluación.

- En general, en el caso de una evaluación que vaya a incluir todas las combinaciones posibles entre n intervenciones, se necesitarán **2^n** grupos.
- Además, para poder distinguir los resultados de los grupos, cada grupo requiere un número suficiente de unidades de observación de modo de garantizar una potencia estadística suficiente. En la práctica, la detección de diferencias entre las ramas de la intervención puede exigir muestras más grandes que la comparación entre un grupo de tratamiento y un grupo de comparación puro. Si las dos ramas de tratamiento logran provocar cambios en los resultados deseados, se requerirán muestras más grandes para detectar las posibles diferencias menores entre los dos grupos.

Probar el impacto de múltiples intervenciones también tiene una implicación más sutil: *a medida que se incrementa el número de intervenciones o niveles de tratamiento que se contrastan unos con otros, se aumenta la probabilidad de encontrar un impacto en al menos una de las pruebas, aunque no haya impacto.* En otras palabras, hay más probabilidades de encontrar un *falso positivo*. Para evitar esto, se deben ajustar las pruebas estadísticas de modo de dar cuenta de las pruebas de hipótesis múltiples. *Los falsos positivos también se denominan errores de tipo II.*

Por último, los *diseños cruzados* también se pueden utilizar en diseños de evaluación que combinan diversos métodos de evaluación. Las reglas operativas que rigen la asignación de cada tratamiento determinarán qué combinación de métodos debe usarse. Por ejemplo, puede ocurrir que el primer tratamiento se asigne sobre la base de una puntuación de elegibilidad, pero el segundo se asignará de manera aleatoria. En este caso, el diseño puede recurrir a un diseño de *regresión discontinua* para la primer intervención y a un método de *asignación aleatoria* para la segunda intervención.

CAPÍTULO 9.

Elección de la Muestra y Diseño de Recolección de Datos

Hasta este momento ya se ha elegido el método para seleccionar el grupo de comparación que sirve de base para estimar el contrafactual, por lo que la siguiente etapa, consiste en determinar qué datos se requieren y la muestra de la que se estimará con precisión, las diferencias entre el grupo de tratamiento y el grupo de comparación. Es importante considerar los criterios sobre el análisis de cómo extraer una muestra de una población de interés (muestreo) y cómo determinar el tamaño que debe tener la muestra para declararla representativa que sirvan de base para generar estimados del impacto del programa (*cálculos de potencia*). El muestreo y los *cálculos de potencia* requieren de ciertas habilidades técnicas específicas que suelen ser encargados a expertos afines, por lo que debe haber una comunicación contínua entre el investigador y el analista de los datos en el que ambos esten de acuerdo con el muestreo y cálculo de potencia, en donde se destacan los elementos que los responsables de la introducción de innovaciones requieren proveer a los expertos en datos.

Determinando la muestra

Todo comienza conel muestreo, el cual se define como: *el proceso de extraer unidades de una población de interés para estimar las características de la población* y es necesario, dado que, frecuentementre, no es posible observar y medir directamente los resultados para toda la población de interés. Los principios de muestreo se aplican para *extraer muestras representativas*. Se requiere seguir, tres grandes pasos para extraer una muestra:

1. Determinar la población de interés.
2. Definir un marco muestral.
3. Extraer el número de unidades requeridas por los cálculos de potencia del marco muestral.

Por lo tanto, se sugiere seguir:

- Definir, en primer lugar, claramente la población de interés, lo que implica que debe especificar con precisión, la unidad en la población de interés para la cual se medirán los resultados.
- Detalle con claridad la cobertura geográfica o cualquier otro atributo pertinente que caracterice a la población de interés.
- Establezca un marco muestral, como la lista más exhaustiva que se puede obtener de las unidades en la población de interés. Idealmente, debería coincidir exactamente con la población de interés. Un censo actualizado, constituiría un marco muestral ideal. Se suelen utilizar las listas existentes, como los censos de población, los censos de instalaciones o los registros de inscritos. El marco muestral, es base para asegurar que las conclusiones que parten del análisis muestral generalizables para el conjunto de la población. Si el marco muestral no coincide con la población de interés, se crea un *sesgo de cobertura.*
- Cuando se produce un *sesgo de cobertura,* no tienen validez externa los resultados de la muestra para el conjunto de la población de interés sino únicamente para la población incluida en el marco muestral. Depende de la magnitud del sesgo de cobertura, la medida en que las estadísticas calculadas a partir de la muestra, se puedan generalizar a toda la población de interés, es decir: *de la falta de coincidencia entre el marco muestral y la población de interés*. Constituyen un riesgo, ya que los marcos muestrales deben ser rigurosos. Por ejemplo, los datos del censo pueden contener la lista de todas las unidades de producción, pero, si ha transcurrido demasiado tiempo entre el censo y el momento en que se recopilaron los datos de la

muestra, el marco muestral ya no estará actualizado. Los datos del censo, adicionalmente, pueden no contener suficiente información sobre atributos específicos para construir un marco muestral. Por ejemplo, si la población de interés está compuesta por **pbt** creadas a nivel inicial, y el censo no incluye datos sobre su continuación en registro, se requerirían datos complementarios que le restan credibilidad.

Obtenidos el marco muestral y la población de interés, se debe elegir el método que determine la muestra **(Mejía-Trejo, 2019a).**

Métodos de muestreo

Se tienen como los más rigurosos, los métodos de *muestreo probabilístico*, ya que asignan, para cada unidad del marco muestral, una probabilidad bien definida, siendo los principales, los mostrados en la **Tabla 9.1.**

El procedimiento para extraer una muestra, a menudo está determinado por las reglas de elegibilidad del programa que se evalúa, dentro del contexto de una evaluación de impacto. *Si la unidad viable más pequeña de implementación es más grande que la unidad de observación, la asignación aleatoria de los beneficios creará clusters*. Por esta razón, el *muestreo de clústers* aparece a menudo en los estudios de evaluaciones de impacto.

El *muestreo no probabilístico* puede provocar graves errores de muestreo. Por ejemplo, supóngase que se emprende una encuesta nacional pidiendo a un grupo de entrevistadores que recopilen datos de los **pbt** más próximas a las sucursales bancarias en cada pueblo. Al utilizar un procedimiento *de muestreo no probabilístico* de este tipo, es muy probable, que la muestra no sea representativa del conjunto de la población de interés. Lo anterior es porque se produce un *sesgo de cobertura*, dado que las **pbt** remotas no serán estudiadas.

Tabla 9.1. Métodos muestreo probabilístico más utilizados

Método de muestreo	Descripción
Aleatorio	Tienen exactamente la misma probabilidad de ser extraídas,todas las unidades de la población. En ocasiones, se produce una confusión entre el *muestreo aleatorio* y la *asignación aleatoria.*¿Qué sucedería está implementando una evaluación de impacto entrevistando a una muestra aleatoria de participantes y no participantes? Suponga que observa a un grupo de individuos que participan de un programa y a un grupo de individuos que no participan en el programa ¿.qué pasaría si se tomara una muestra aleatoria de cada uno de estos dos grupos? Si los participantes y los no participantes tienen diferentes características, también lo tendrá la muestra de participantes y no participantes. *El muestreo aleatorio no hace que dos grupos no comparables sean comparables y no proporciona validez interna para la evaluación de impacto*. Este es el motivo por el que el muestreo aleatorio no es suficiente para la evaluación de impacto. La asignación aleatoria de los beneficios de un programa es diferente del muestreo aleatorio. El proceso de asignación aleatoria parte de una población de interés elegible y utiliza un procedimiento de aleatorización para asignar las unidades de la población elegible a un grupo de tratamiento que ser objeto de una intervención, y a un grupo de comparación que no lo será. Cuando la asignación aleatoria está bien implementada, contribuye a la validez interna de la evaluación de impacto. El muestreo aleatorio puede ser útil para asegurar la validez externa, en la medida en que la muestra se extrae aleatoriamente de la población de interés.
Aleatorio estratificado	La población se divide en dos grupos y se lleva a cabo un muestreo aleatorio en cada grupo. Por lo tanto, todas las unidades en cada grupo (*o estrato*) tienen la misma probabilidad de ser extraídas. Siempre y cuando todos los grupos sean lo suficientemente grandes, el muestreo estratificado permite formular *inferencias* acerca de los resultados no solo a nivel de la población, sino también dentro de cada grupo. El muestreo estratificado es útil cuando se quiere elaborar una muestra de los subgrupos pequeños en la población (por ejemplo, las minorías) con el fin de estudiarlos más en detalle. Es esencial la *estratificación* para las evaluaciones que buscan comparar los impactos del programa entre esos subgrupos.
Muestreo de clústers	Las unidades se agrupan en clústers (*conglomerados*) y se extrae una muestra aleatoria de los mismos. Posteriormente, o todas las unidades en esos clusters constituyen la muestra, o bien se extrae un cierto número de unidades del clúster de forma aleatoria. Esto significa que cada clúster tiene una probabilidad bien definida de ser seleccionado y las unidades dentro de un clúster seleccionado también tienen una probabilidad bien definida de ser extraídas.

Fuente: Gertler et al. (2017).

En conclusión, es sumamente importante prestar mucha atención tanto al *marco muestral* como al *procedimiento de muestreo* para establecer, si los resultados obtenidos de una determinada muestra, es posible generalizarlos al conjunto de la población de interés. Aun cuando el marco muestral tenga *perfecta cobertura* y se utilice un procedimiento de muestreo probabilístico, *los errores de no muestreo* también pueden afectar la *validez interna y externa* de la evaluación de impacto. Por tanto, se requieren muestras relativamente más grandes.

- Para obtener estimaciones precisas de las características de la población.
- Para poder obtener estimaciones precisas de las diferencias entre grupos de tratamiento y de comparación, es decir, para estimar el impacto de un programa.

Tabla 9.1. Métodos muestreo probabilístico más utilizados

Método de muestreo	Descripción
Aleatorio	Tienen exactamente la misma probabilidad de ser extraídas,todas las unidades de la población. En ocasiones, se produce una confusión entre el *muestreo aleatorio* y la *asignación aleatoria*. ¿Qué sucedería está implementando una evaluación de impacto entrevistando a una muestra aleatoria de participantes y no participantes? Suponga que observa a un grupo de individuos que participan de un programa y a un grupo de individuos que no participan en el programa ¿qué pasaría si se tomara una muestra aleatoria de cada uno de estos dos grupos? Si los participantes y los no participantes tienen diferentes características, también lo tendrá la muestra de participantes y no participantes. *El muestreo aleatorio no hace que dos grupos no comparables sean comparables y no proporciona validez interna para la evaluación de impacto*. Este es el motivo por el que el muestreo aleatorio no es suficiente para la evaluación de impacto. La asignación aleatoria de los beneficios de un programa es diferente del muestreo aleatorio. El proceso de asignación aleatoria parte de una población de interés elegible y utiliza un procedimiento de aleatorización para asignar las unidades de la población elegible a un grupo de tratamiento que ser objeto de una intervención, y a un grupo de comparación que no lo será. Cuando la asignación aleatoria está bien implementada, contribuye a la validez interna de la evaluación de impacto. El muestreo aleatorio puede ser útil para asegurar la validez externa, en la medida en que la muestra se extrae aleatoriamente de la población de interés.
Aleatorio estratificado	La población se divide en dos grupos y se lleva a cabo un muestreo aleatorio en cada grupo. Por lo tanto, todas las unidades en cada grupo (*o estrato*) tienen la misma probabilidad de ser extraídas. Siempre y cuando todos los grupos sean lo suficientemente grandes, el muestreo estratificado permite formular *inferencias* acerca de los resultados no solo a nivel de la población, sino también dentro de cada grupo. El muestreo estratificado es útil cuando se quiere elaborar una muestra de los subgrupos pequeños en la población (por ejemplo, las minorías) con el fin de estudiarlos más en detalle. Es esencial la *estratificación* para las evaluaciones que buscan comparar los impactos del programa entre esos subgrupos.
Muestreo de clusters	Las unidades se agrupan en clústers (*conglomerados*) y se extrae una muestra aleatoria de los mismos. Posteriormente, o todas las unidades en esos clústers constituyen la muestra, o bien se extrae un cierto número de unidades del clúster de forma aleatoria. Esto significa que cada clúster tiene una probabilidad bien definida de ser seleccionado y las unidades dentro de un clúster seleccionado también tienen una probabilidad bien definida de ser extraídas.

Fuente: Gertler et al. (2017).

Cálculo de potencia de muestra

A fin de estimar las características de una población, el muestreo describe el proceso para elaborar una muestra de unidades de esa población de interés. Las muestras más grandes dan estimaciones más precisas de las características de la población, pero ¿de qué tamaño exactamente, tienen que ser las muestras para una evaluación de impacto?

Los cálculos para determinar el tamaño de la muestra, se denominan cálculos de potencia. El caso más sencillo, a saber: una evaluación realizada utilizando un *método de asignación aleatoria*, para probar la efectividad de un programa en relación con un grupo de comparación que no recibe una intervención, y suponiendo que el incumplimiento no es un problema. Calcular la potencia es calcular el tamaño mínimo de la muestra que es necesario para llevar a cabo una evaluación de impacto y para responder de forma convincente a la pregunta de interés para las políticas de introducción de innovaciones. Ver **Tabla 9.2.**

Tabla 9.2. Razones para realizar cálculos de potencia

- Evaluar si las bases de datos existentes, son suficientemente grandes para llevar a cabo una evaluación de impacto.
- Evitar recopilar pocos datos. Si la muestra es demasiado pequeña, es posible que no detecte un impacto positivo, aunque existiera y, por lo tanto, también es posible llegar a la conclusión de que no ha tenido efecto. Esto podría provocar una decisión de política de introducción de innovaciones tendientes a eliminar el programa, lo cual sería perjudicial.
- Contribuir a tomar decisiones a propósito del tamaño adecuado de la muestra. Los tamaños más grandes de la muestra proporcionan estimaciones más precisas de los impactos del programa, pero la recopilación de información puede ser muy onerosa. Los cálculos de potencia proporcionan insumos clave para evaluar el equilibrio entre los costos requeridos para recopilar más datos y los beneficios de una mayor precisión en la evaluación de impacto.

Fuente: Gertler et al. (2017).

La muestra más pequeña que permitirá detectar diferencias significativas en los resultados entre los grupos de tratamiento y comparación, lo hace el cálculo de potencia; constituyen una indicación de la muestra más pequeña (y el presupuesto más bajo) con el que es posible medir el impacto de un programa. Los cálculos de potencia, tienen la capacidad de determinar cuáles son los programas que tienen éxito y cuáles no.

La pregunta básica de la evaluación de impacto es: ¿cuál es el impacto o efecto causal de un programa en un resultado de interés? o ¿el impacto del programa es diferente de cero? En el caso de la asignación aleatoria, es posible responderla en:

- Estimar los resultados promedio para los grupos de tratamiento y comparación.
- Valorar si existe una diferencia entre el resultado promedio del grupo de tratamiento y el resultado promedio del grupo de comparación.

Grupo tratamiento vs. comparación

Suponga que se debe estimar el impacto de un programa de estímulos a la innovación (**pei**) nacional de pymes de base tecnológica (**pbt**) en un innovador proceso de riego-fertilización-insecticida **100%** natural para producir maíz por hectárea, a los **2** años de inciado, y que hay **200 pbt** elegibles para el programa. Del total de **pbt** elegibles, **100** fueron asignados de forma aleatoria para participar en el programa. Los **100 pbt** elegibles que no fueron asignados aleatoriamente al programa sirven como grupo de comparación, así, se tiene:

- Primer paso, estimar la producción promedio de maíz por hectárea reportado por las **pbt** que participaron y de los que no participaron. Para determinar la producción promedio de maíz por hectárea de las **pbt** que participaron, se podría recopilar el **peso** de la producción de maíz por hectárea, de cada una de las **100 pbt** y luego calcular el promedio. Desde luego, sería un procedimiento sumamente costoso.

- Afortunadamente, no es necesario verificar el peso de la producción de maíz por hectárea de cada **pbt**. El promedio se puede estimar utilizando el peso promedio de maíz por hectárea de una muestra extraída de la población de los **pbt** que participan. Cuantos más **pbt** haya en la muestra, más cerca estará el promedio estimado del promedio real.
- Cuando una muestra es pequeña, el peso promedio de maíz por hectárea constituye una estimación muy imprecisa del promedio en la población. Por ejemplo, una muestra de peso de maíz por hectárea producido por la **pbt** no dará una estimación precisa. En cambio, una muestra de peso de maíz por hectárea de **100 pbt,** producirá una estimación más precisa mucho más cercana al verdadero peso promedio de maíz por hectárea. En general, cuantas más observaciones haya en la muestra, más precisas serán las estadísticas obtenidas de la muestra.
- Por lo tanto, se sabe que con una muestra más grande de peso de maíz por hectárea, se obtendrá una imagen más exacta de la población de las **pbt** que participan. Lo mismo ocurrirá con las **pbt** que no participan: a medida que crece el tamaño de la muestra de peso de maíz por hectárea de estos últimos, se sabe con mayor precisión cómo es esa población. ¿Pero por qué habría esto de importar? Si se puede estimar el resultado promedio (el peso de maíz por hectárea) de la producción por **pbt** que participan y no participan con más precisión, también se podrá saber con más precisión la diferencia de pesos muestra de maíz por hectáreas producido entre ambos grupos, y eso es el impacto del programa.

Dicho de otra manera, si solo se tiene una idea vaga del peso promedio del maíz por hectárea producido por las **pbt** en los grupos que participan (tratamiento) y que no participan (comparación), ¿cómo se podrá tener una idea precisa de la diferencia de peso de los dos grupos? La verdad es que no es posible.

Grupo tratamiento vs. comparación y sus resultados

Una vez estimado el resultado promedio del peso de maíz por hectárea producido del grupo de tratamiento (**pbt** que participan seleccionados por asignación aleatoria) y el grupo de comparación (los **pbt** que no participan seleccionados por asignación aleatoria), es posible proceder a determinar si los dos resultados son diferentes. Debe ser claro que el *procedimiento básico es restar los promedios y calcular la diferencia.* En la evaluación de impacto y a nivel estadístico, se pone a prueba la *hipótesis nula vs. la hipótesis alternativa.*

La hipótesis nula, es la hipótesis de que el programa no tiene un impacto.

H0: impacto o diferencia entre el resultado en el grupo de tratamiento y comparación = 0.

Ha: impacto o diferencia entre el resultado en el grupo de tratamiento y comparación ≠ 0.

En el ejemplo del programa de estímulos a la innovación (**pei**) nacional de pymes de base tecnológica (**pbt**) en un innovador proceso de riego-fertilización-insecticida **100%** natural para producción de maíz por hectárea, se comienza con una muestra del peso de producción por hectárea de **2 pbt:** tratado y de comparación. Con una muestra tan pequeña, la estimación del peso de maíz por hectárea promedio de los **pbt** tratados y los **pbt** de comparación y, por lo tanto, la estimación de la diferencia entre los dos grupos, no será fiable. Es posible verificar esto extrayendo diferentes muestras de peso de maíz por hectárea de dos **pbt** del grupo de tratamiento y dos **pbt** del grupo de comparación. Lo que se encontrará es que el impacto estimado del programa tiene una alta variación.

Suponga por el contrario, que comienza con una muestra de pesos de maíz por hectárea de **10 pbt** tratados y **10 pbt** del grupo de comparación. Como se señaló, las estimaciones del peso promedio de ambos grupos serán

mucho más precisas. Por lo tanto, la estimación de la diferencia entre los dos grupos también lo será.

Ahora bien, suponga que se observa que el peso promedio en la muestra de los pesos de maíz por hectárea de **pbt** del tratamiento (que participan) y es de **12.2** ton/ha, y el promedio de los niños en la muestra de comparación (que no participan) es de **12.0** ton/ha. La diferencia entre ambos grupos es de **0.2** ton/ha. Si estas cifras correspondieran a muestras de dos observaciones cada una, no se sabría bien si el impacto del programa es verdaderamente positivo porque esos **0.2** ton/ha podrían deberse a la falta de precisión en las estimaciones. Sin embargo, si estas cifras provienen de muestras de **1,000** observaciones cada una, aumentaría la confianza de que se acercan bastante al verdadero impacto del programa, que en este caso sería positivo.

Por lo tanto, se debe establecer como pregunta clave por el investigador:

¿Cuál es exactamente el tamaño que debe tener la muestra para que nos permita determinar que un impacto estimado positivo se debe al verdadero impacto del programa y no a una falta de precisión en las estimaciones?

Errores potenciales que se deben evitar

Los tipos de errores estadísticos típicos que se tienden a presentar, se muestran en la **Tabla 9.3.**

Tabla 9.3. Errores estadísticos

Nivel de Error	Descripción
I o Alfa	Probabilidad de rechazar correctamente la hipótesis nula cuando es cierta (falso *positivo*). *Se comete si una evaluación concluye que el programa ha tenido impacto, cuando en realidad no lo ha tenido.* En el caso **pei** de las **pbt** de proceso de innovación riego-fertilización-insecticida **100%** natural para producir maíz, esto ocurriría si el equipo de evaluación, concluyera que el peso de maíz por hectárea promedio producido por las **pbt** de la muestra tratada es superior al de las **pbt** de la muestra de comparación, aunque el peso del maíz por hectárea promedio de los **pbt** de dos poblaciones fuera, de hecho, igual y las diferencias observadas fueran solo coincidencia. En este caso, el impacto positivo que se observara provendría únicamente de la falta de precisión de las estimaciones. Al probar la hipótesis de que un programa ha tenido impacto, los estadísticos pueden limitar el tamaño de los errores de tipo I. La probabilidad de un error de tipo I se establece mediante un el *nivel de significancia,* el cual, suele fijarse en **5%**, lo que quiere decir que se puede tener un **95%** de confianza en llegar a la conclusión de que el programa ha tenido impacto. Si se es más rigosrista de no ometer un error de tipo I, es posible establecer un nivel de significancia menor, por ejemplo, del **1%,** de manera que tenga un **99%** de confianza de llegar a la conclusión de que el programa ha tenido impacto.

II o Beta	Probabilidad rechazar incorrectamente la hipótesis nula cuando es realmente falsa. Es el tipo contrario de error (*falso negativo*). *Se comete cuando una evaluación llega a la conclusión de que el programa no ha tenido impacto, cuando en realidad sí lo ha tenido*. En el caso **pei** de las **pbt** de proceso de innovación riego-fertilización-insecticida **100%** natural para producir maíz, esto ocurriría si se concluyera quel peso de maíz por hectárea promedio producido por las **pbt** en las dos muestras es el mismo, aunque el peso de maíz por hectárea promedio producido de la población de tratamiento es, de hecho, superior al de los pbt del grupo de comparación. Una vez más, el impacto debería haber sido positivo, pero debido a la falta de precisión de las estimaciones, se llega a la conclusión de que el programa ha tenido un impacto cero. Este tipo de errores también debe preocupar a los responsables de las políticas de introducción de innovaciones ya que influyen numerosos factores en la probabilidad de cometer un error de tipo II, pero el tamaño de la muestra es crucial. Si el peso de maíz por hectárea promedio de **50,000 pbt** tratados es el mismo que el peso de maíz por hectárea promedio de **50.000 pbt** de comparación, es probable que se concluya que el programa no ha tenido impacto. Al contrario, si en una muestra de dos **pbt** del grupo de tratamiento estos tienen un peso de maíz por hectárea promedio igual que en el caso de la muestra de dos **pbt** del grupo de comparación, es más difícil llegar a una conclusión fiable. Así ¿el peso de maíz por hectárea promedio es similar porque la intervención ha tenido impacto o porque los datos no son suficientes para comprobar la hipótesis en una muestra tan pequeña? Las muestras grandes reducen la probabilidad de que solo se observe a los **pbt** con peso de maíz por hectárea promedio que sean iguales por una cuestión casual o de *mala suerte*. En las muestras grandes, la diferencia de promedios entre la muestra tratada y la muestra de comparación proporciona una mejor estimación de la verdadera diferencia de los promedios entre todas las unidades tratadas y todas las unidades de comparación.
Potencia de una Prueba 1-Beta	Probabilidad de rechazar correctamente la hipótesis nula cuando debe ser rechazada. No se plantean niveles aceptables tanto de alfa como de beta, debido a que los errores de Tipo I y Tipo II están inversamente relacionados, esto es que a medida que el error tipo Alfa se hace más restrictivo (cercano a cero), el error tipo B aumenta. Al disminuir el error de Tipo I también se reduce el poder de la prueba estadística. La potencia de una prueba es igual a 1 menos la probabilidad de un error de tipo II.

Fuente: Gertler et al. (2017) y Mejía-Trejo (2019a).

Cálculo de potencia estadística

En una evaluación de impacto, la *potencia* (o *potencia estadística*), es la probabilidad de detectar una diferencia entre los grupos de tratamiento y comparación cuando esta, de hecho existe. Así, existe una alta potencia si hay un bajo riesgo de no detectar verdaderos impactos del programa, es decir, de cometer un *error de tipo II* en una evaluación de impacto de programa. Como se aprecia, el tamaño de la muestra es un factor determinante crucial de la potencia en una evaluación de impacto. Determinar el tamaño de una muestra para evitar llegar a la conclusión de que un programa no ha tenido impacto, cuando de hecho sí lo ha tenido (*error de tipo II*), es el objetivo central del cálculo de potencia La potencia de una prueba es igual a 1 menos la probabilidad de un *error de tipo II*.

Una potencia elevada se tiene, en una evaluación de impacto de programa, si es poco probable que se produzca un *error de tipo II*, lo que significa que es poco probable la investigación demerite por los resultados que muestren que el programa que se evalúa no ha tenido impacto, cuando en realidad sí lo ha tenido.

Desde una perspectiva de políticas de introducción de innovaciones, las evaluaciones de impacto:

- Con insuficiente potencia, con una alta probabilidad de *errores de tipo II*, no solo son inútiles sino que también pueden resultar muy costosas.
- Una alta probabilidad de un *error de tipo II* pone en peligro el potencial de una evaluación de impacto de identificar resultados estadísticamente significativos.
- Por lo tanto, destinar recursos a evaluaciones de impacto sin suficiente potencia es una inversión riesgosa.
- Las evaluaciones de impacto sin suficiente potencia también pueden tener graves consecuencias prácticas. Por ejemplo, en la intervención hipotética del programa de estímulos a la innovación (**pei**) nacional de pymes de base tecnológica (**pbt**) en un innovador proceso de

riego-fertilización-insecticida **100%** natural para producción de maíz por hectárea mencionada, si se llegara a la conclusión de que el programa no fue efectivo, aunque sí lo fue, los responsables de las políticas de introducción de innovaciones podrían poner fin a un programa que, de hecho, beneficia a las **pbt** y al público consumidor. Por lo tanto, *es crucial minimizar la probabilidad de errores de tipo II utilizando muestras lo suficientemente grandes en las evaluaciones de impacto*. Por esto es tan fundamental y pertinente llevar a cabo cálculos de potencia.

Para realizar cálculos de potencia de un programa de asignación aleatoria, se sugiere responder las preguntas, que dependen del contexto y momento en que se generan al momento de las políticas de introducción de innovaciones. **Ver Tabla 9.4.**

Tabla 9.4. Preguntas que esclarecen el cálculo de potencia de muestra y estadística

P1. ¿El programa funciona mediante clústers?
P2. ¿Cuál (es) es/son los indicadores de resultados?
P3. ¿Cuál es el nivel mínimo de impacto que justificaría la inversión hecha en la intervención?
P4. ¿Cuál es la media de resultado para la población de interés? ¿Cuál es la varianza subyacente del indicador de resultado?
P5. ¿Cuáles son los niveles razonables de potencia estadística y de significancia e estadística en la evaluación que se llevará a cabo?

Fuente: Gertler et al. (2017) con adaptación propia.

Respondiendo la necesidad del funcionamiento con clúster

El primer paso, en los cálculos de potencia, consiste en determinar si el programa de introducción de innovaciones que se quiere evaluar, genera *clusters* a lo largo de su implementación. Se generan clústers en torno al lugar de una intervención cuyo nivel (a menudo, lugares) es diferente del nivel al que se querrían medir los resultados (a menudo, personas), sin embargo el impacto se mide en los sujetos del clúster. Cuando una evaluación de impacto genera clústers, es el número de estos lo que determina en gran parte el tamaño de la muestra útil. En cambio, el número de individuos en los clústers importa menos.

La naturaleza de cualquier dato de la muestra construido a partir de programas que están conglomerados ,es algo diferente de las muestras obtenidas a partir de programas que no lo están. Por tanto, los cálculos de potencia tienen cierta diferencia (ligera) en su pasos de ejecución, dependiendo de si un programa asigna aleatoriamente los beneficios entre los clústers o sencillamente asigna los beneficios aleatoriamente entre todas las unidades de una población.

Para lograr un mejor entendimiento, contrastamos el funcionamiento sin/con clúster mostrados en la **Tabla 9.5.**

Tabla 9.5. Clúster (ausencia/presencia)

Solución **P1**: Ausencia Clúster El programa que se evalúa asigna de forma aleatoria los beneficios entre todas las unidades de la población elegible.	Solución **P1**: Presencia Clúster Algunos programas asignan beneficiosa nivel de clústers. Un principio rector clave es que el número de clústers suele importar mucho más que el número de individuos en los clústers.
Solución **P2.** Estos indicadores derivan del objetivo del programa, de una teoría del cambio y de la pregunta fundamental de la investigación de la evaluación, como se señaló en la primera parte. Los cálculos de potencia también ayudarán a entender el tipo de indicadores más adecuados para las evaluaciones de impacto. En realidad, es posible se requieran muestras de diversos tamaños para medir impactos en diferentes indicadores. Solución **P3.** Se debe determinar el impacto mínimo que justificaría la inversión realizada en la intervención. Se trata sobre todo de una pregunta de políticas de introducción de innovaciones, más que de una pregunta técnica. La respuesta es sumamente específica del contexto, pero en todos los casos es necesario determinar el cambio en los indicadores de resultados que justificaría la inversión hecha en el programa. Dicho de otra manera, ¿cuál es el nivel de impacto por debajo del cual una intervención debería considerarse no exitosa? Esto es el *efecto mínimo detectable* **(EMD)** que la evaluación de impacto tiene que ser capaz de identificar. Esto depende no solo del costo del programa y del tipo de beneficios que proporciona, sino también del costo de oportunidad de no invertir fondos en una intervención alternativa. Si bien los **EMD** tienden a basarse en objetivos de las políticas de introducción de innovaciones, es posible utilizar otros enfoques para establecerlos. Posiblemente sea útil tomar como referencia **EMD** en relación con resultados de los estudios en programas similares que den pistas sobre la magnitud de los impactos que se deban esperar.	Se requiere un número suficiente de clusters para probar de forma convincente si un programa ha tenido impacto al contraponer resultados en muestras de las unidades de tratamiento y comparación. Es el número de clusters el que determina en gran parte el tamaño de la muestra útil o efectivo. *Si se asigna de manera aleatoria el tratamiento entre un pequeño número de clústers, es poco probable que los clústers de tratamiento y comparación sean idénticos.* La asignación aleatoria entre dos **pbt**, no garantizará que los dos clústers sean similares. En cambio, la asignación aleatoria de una intervención entre **100 pbt**, tiene más probabilidades de asegurar que los grupos de tratamiento y comparación sean similares. En resumen, se requiere un número suficiente de clústers para alcanzar un equilibrio. Además, el número de clústers también importa para la precisión de los efectos estimados del tratamiento. Se requiere un número suficiente de clústers para probar la hipótesis de que un programa tiene un impacto con suficiente potencia. Cuando se implementa una evaluación de impacto basada en la asignación aleatoria, es muy importante asegurar que el número de clústers sea suficientemente grande. Se puede establecer el número de clústers requeridos para pruebas de hipótesis precisas efectuando cálculos de potencia. Esto exige formular las mismas cinco preguntas expuestas anteriormente, más una pregunta extra:

Por ejemplo, determinar qué % de desviaciones del programa son pequeñas o grandes mediante simulaciones ex ante normalmente económicos (tir, gastos, etc.). Es más fácil identificar una gran diferencia entre dos grupos que identificar una diferencia pequeña. En una de éstas últimas, se necesitará una estimación muy precisa de la diferencia de los resultados medios entre los dos grupos. Esto requiere una muestra grande. Como alternativa, en las intervenciones que se consideran viables solo si generan grandes cambios en los indicadores de resultado, las muestras necesarias para llevar a cabo una evaluación de impacto, serán más pequeñas. Sin embargo, el **EMD** debería fijarse de manera conservadora, dado que es menos probable que se detecte cualquier impacto menor que el **EMD.** Solución **P4.** Para llevar a cabo cálculos de potencia, se deben estimar algunos parámetros básicos, como el *promedio de la línea de base* y una *varianza de los indicadores de resultado*. Estos valores de referencia deberían preferiblemente obtenerse de los datos recopilados en un contexto similar a aquel en el cual se implementará el programa que se estudia, o de una encuesta piloto en la población de interés. Cuanto más variables sean los resultados de interés, mayor será la muestra que se necesitará para estimar un efecto de tratamiento preciso. En el ejemplo de las **pbt**, el peso de maíz por hectárea promedioes el resultado de interés. Si todos los individuos pesan lo mismo en la *línea de base*, será factible estimar el impacto de una intervención en el **pei** en una muestra pequeña. En cambio, si los pesos de *línea de base* de los **pbt** son muy variables, se requerirá una muestra más grande para estimar el impacto del programa.	*¿Cuán variable es el indicador de resultado en los clústers?* En el extremo, todos los resultados en un cluster están perfectamente correlacionados. Por ejemplo, puede ocurrir que el ingreso económico de un **pbt** no varíe especialmente en las comunidades, pero que entre comunidades se observe una desigualdad importante en los ingresos. En este caso, si usted considera añadirlo a su muestra de evaluación, agregar a un individuo de una comunidad nueva aumentará mucho más la potencia que introducir un individuo de una comunidad que ya está representada. Dado que los resultados están plenamente correlacionados en un clúster, añadir un nuevo individuo de ese clúster existente no aportará nueva información. En realidad, en este caso, es probable que el individuo de la segunda comunidad tenga un aspecto muy similar al individuo original ya incluido. En general, una mayor correlación intra-clúster en los resultados (es decir, una mayor correlación en los resultados o características entre las unidades que pertenecen al mismo clúster) aumenta el número de clústers requeridos para alcanzar un *determinado nivel de potencia*. En las muestras con clústers, los cálculos de potencia subrayan los beneficios relativos entre añadir clústers y añadir observaciones dentro de los clústers. El aumento relativo de la potencia al agregar una unidad de un nuevo clúster es casi siempre mayor que el de sumar una unidad a un clúster ya existente. Aunque el incremento de la potencia al añadir un nuevo cluster puede ser drástico, agregar clústers también puede tener implicaciones operativas y elevar el costo de la implementación del programa o de la recopilación de datos.

Solución **P5.** El equipo de evaluación tiene que determinar un nivel de potencia razonable y un nivel de significancia para la evaluación de impacto planificada. Como ya se señaló, la potencia de una prueba es igual a **1** menos la probabilidad de cualquier *error de tipo II*. Por lo tanto, la potencia oscila entre **0** y **1**, donde un valor alto indica menos riesgo de no identificar un impacto existente. *Una potencia de **0.8** es una referencia generalmente utilizada para los cálculos de potencia*. Significa que se encontrará un impacto en el **80%** de los casos allí donde se haya producido. Un nivel más alto de potencia de **0.9** (o **90%**) a menudo proporciona una referencia útil pero más conservadora, lo cual aumenta el tamaño requerido de la muestra. *El nivel de significancia es la probabilidad de cometer un error de tipo I.* Normalmente se fija en **5%,** de modo que se puede tener una confianzadel **95%** de llegar a la conclusión de que el programa ha tenido impacto si se encuentra un impacto significativo. Otros niveles habituales de significancia son **1%** y **10%.** *Cuanto menor sea el nivel de significancia, más confianza se puede tener en que el impacto estimado es real.*	En numerosos casos, se requieren al menos entre **40** y **50** clusters en cada grupo de tratamiento y comparación para obtener potencia suficiente y garantizar la similitud de las características de línea de base al usar métodos de asignación aleatoria. Sin embargo, es posible que el número varíe de acuerdo con los diversos parámetros ya analizados, así como la correlación *intra-clúster*. Además, es probable que el número probablemente aumente al utilizar métodos distintos de la asignación aleatoria (suponiendo que todos los demás factores permanezcan constantes).

Fuente: Gertler et al. (2017) con adaptación propia.

Una vez que se han abordado las cinco preguntas, se procede a calcular el tamaño requerido de la muestra utilizando software estadístico. El cálculo de potencia indicará el tamaño requerido de la muestra, dependiendo de los parámetros establecidos en las preguntas del 1 al 5. Los propios cálculos son sencillos, una vez que se han determinado los parámetros relevantes para las políticas de introducción de innovaciones. Se siugiere incluir un análisis de la sensibilidad del cálculo de potencia ante cambios en los supuestos. En otras palabras, es importante entender cuánto tendrá que aumentar el tamaño requerido de la muestra con supuestos más conservadores (como un impacto previsto menor, mayor varianza en el indicador de resultado o un mayor nivel de potencia). Encargar cálculos de potencia para diversos indi-

cadores de resultados, es también una buena práctica dado que los tamaños requeridos de la muestra pueden variar considerablemente si algunos indicadores de resultados son mucho más variables que otros. Los cálculos de potencia también pueden indicar el tamaño de la muestra necesario para establecer una comparación de los impactos del programa en diferentes subgrupos específicos. Cada subgrupo tendría que tener el tamaño requerido de la muestra.

Más allá del clúster y la asignación aleatoria

Como se mencionó, una evaluación de impacto basada en clúster con asignación aleatoria, el escenario más sencillo y, por lo tanto, el más adecuado para transmitir la intuición en que se basan los cálculos de potencia. Sin embargo, existen aún numerosos aspectos prácticos de los cálculos de potencia que aún no se han sido analizados, y es necesario considerar detenidamente las desviaciones de los casos básicos que se abordan aquí. Ver la **Tabla 9.6.**

Tabla 9.6. Otros escenariuos y técnicas a tomar en cuenta

Utilización de métodos cuasi experimentales. *Ceteris paribus*
Los métodos de evaluación de impacto cuasi experimentales, como la **RD**, **PMS** o las **DD** tienden a requerir muestras más grandes que el método de referencia de asignación aleatoria. Por ejemplo, al utilizar el diseño de **RD (capítulo 7)** se subrayó que solo se pueden considerar las observaciones en torno al *umbral de elegibilidad*. Se requiere una muestra suficientemente grande en torno a ese umbral. Los cálculos de potencia son necesarios para estimar la muestra requerida de modo de establecer comparaciones significativas en torno al umbral. Por otro lado, la disponibilidad de diversas rondas de datos puede contribuir a aumentar la potencia de una evaluación de impacto con un determinado tamaño de la muestra. Por ejemplo, los *datos de línea de base* sobre resultados y otras características pueden añadir precisión a la estimación de los efectos de tratamiento. La disponibilidad de medidas repetidas de resultados después del comienzo del tratamiento también puede ser útil.

Análisis de diferentes modalidades de programa o innovaciones de diseño

En los ejemplos presentados en este capítulo, el tamaño total de la muestra se dividía por igual entre los grupos de *tratamiento* y *comparación*. En algunos casos, la principal pregunta de las políticas de introducción de innovaciones con respecto a la evaluación, puede generar la comparación de impactos del programa entre las modalidades del programa o las innovaciones de diseño. Si esto es así, el impacto previsto puede ser relativamente menor que si un grupo de tratamiento objeto de un programa fuera comparado con un grupo de comparación que no recibía ningún tipo de beneficios. Como tal, el **EMD** entre los dos grupos de tratamiento puede ser más pequeño que el **EMD** entre el grupo de tratamiento y el grupo de comparación. Esto implicaría que la distribución óptima de la muestra generaría grupos de tratamiento que son relativamente más grandes que el grupo de comparación. En las evaluaciones de impacto con múltiples ramas de tratamiento, puede que sea necesario implementar cálculos de potencia para estimar por separado el tamaño de cada grupo de tratamiento y comparación, en función de la principal pregunta de interés de las políticas de introducción de innovaciones.

Comparación de subgrupos

En otros casos, algunas de las preguntas de la evaluación de impacto pueden centrarse en estimar si los impactos de un programa varían entre diferentes subgrupos, como el género, la educación, la edad o las categorías de ingreso. Si esto es lo que ocurre, *los requisitos del tamaño de la muestra serán mayores y los cálculos de potencia tendrán que ajustarse de forma correspondiente*. Por ejemplo, una pregunta clave de las políticas de inytroducción de innovaciones puede ser si un programa como el **pei** opara **pbt** tiene un impacto mayor en los CEO mujer que en los CEO hombre. Se necesitará un número suficiente de CEOs de cada género en el grupo de tratamiento y el grupo de comparación para detectar un impacto en cada subgrupo. Si se pretende comparar los impactos del programa entre dos subgrupos, es posible que se duplique el tamaño requerido de la muestra. Si se considera la heterogeneidad entre más grupos (por ejemplo, por la edad) también puede aumentar considerablemente el tamaño requerido de la muestra. Si este tipo de comparaciones entre grupos ha de llevarse a cabo en el contexto de una evaluación de impacto que depende de la *asignación aleatoria*, es preferible también tenerlas en cuenta cuando se implementa la aleatorización y, sobre todo, para aplicar una asignación aleatoria *por bloques o estratos* (es decir, en cada subgrupo que se compara). En la práctica, aunque no se realice ninguna comparación entre subgrupos, la *aleatorización estratificada o por bloque* puede contribuir a maximizar aún más la potencia de un determinado tamaño de la muestra.

Análisis de múltiples resultados
Es necesario proceder con singular cuidado cuando se emprenden cálculos de potencia en los casos, en que una evaluación de impacto pretenda probar si un programa genera cambios en múltiples resultados. Si se tienen en cuenta numerosos resultados diferentes, habrá una probabilidad relativamente más alta de que la evaluación de impacto, encuentre impactos en uno de los resultados solo por azar. Para abordar esto, el equipo de evaluación de impacto tendrá que pensar en probar la *significancia estadística conjunta* de los cambios en diversos resultados. Como alternativa, se pueden elaborar algunos índices o familias de resultados. Estos enfoques para lidiar con las pruebas de múltiples hipótesis tienen implicaciones para los cálculos de potencia y el tamaño de la muestra y, en ese sentido, hay que tenerlos en cuenta cuando se define la muestra necesaria para la evaluación de impacto.
Contra el cumplimiento imperfecto o el desgaste de la muestra
Los cálculos de potencia, suelen proporcionar el tamaño mínimo requerido de la muestra. *En la práctica, los problemas de implementación a menudo implican que el tamaño de la muestra real es más pequeño que el tamaño planificado*. Por ejemplo, el cumplimiento imperfecto puede significar que solo se inscribe una parte de los beneficiarios a los que se ofrece el programa. Los requisitos del tamaño de la muestra aumentan cuando surge el cumplimiento imperfecto. Además, aunque todos los individuos se inscribieran en el programa, es posible se produzca algún grado de desgaste en la encuesta de seguimiento si no se da con el paradero de todos los individuos. Aunque ese incumplimiento o desgaste es aleatorio y no afecta la consistencia de las estimaciones de impacto, *estos aspectos influirían en la potencia de la evaluación de impacto*. Para dar cuenta de dichos factores, generalmente se recomienda *añadir un margen al tamaño de la muestra prevista por los cálculos de potencia*. De la misma manera, los datos de menor calidad tendrán más error de medición y harán que los resultados de interés sean más variables, además de que requerirán tamaños de la muestra más grandes.

Fuente: Gertler et al. (2017) con adaptación propia.

Como se apreciará, las reflexiones más avanzadas mencionadas en esta sección exceden el alcance de este libro. En la práctica, se sugiere la integración de los equipos de evaluación que deban incluir o contratar a expertos que efectúen cálculos de potencia y asesoramiento en temas más complejos estadísticos.

Recomendaciones para recopilar datos

En toda evaluación de impacto, los datos existentes disponibles y declarados suficientes normalmente son raros, por lo que muy posiblemente habrá que depender de datos de encuestas **(Mejía-Trejo, 2019b)** y presupuestar la recopilación de datos nuevos. Aunque esto suele implicar el mayor costo de una evaluación de impacto, también proporciona la flexibilidad para garantizar que se midan todos los indicadores necesarios en una evaluación integral del desempeño del programa. Las encuestas son:

- La base de recopliación de datos de la gran mayoría de las evaluaciones de impacto,
- Incluyen al menos una *encuesta de línea de base* antes de la intervención o innovación que se evalúa, y
- Una encuesta de seguimiento después de que se ha implementado la intervención.

En función del programa que se evalúa y de la unidad de análisis, los datos de las encuestas pueden ser de diversos tipos, Por ejemplo:

- Las encuestas de empresas utilizan a las firmas como la principal unidad de observación,
- Las encuestas de instalaciones, utilizan las pymes, centros de salud o las escuelas como la principal unidad de observación, y
- Las encuestas de hogares utilizan los hogares como la principal unidad de observación.

La mayoría de las evaluaciones dependen de encuestas individuales o de organizaciones como pymes, hogares como fuente primaria de datos. Se revisan algunos principios generales de la recopilación de datos de las encuestas, estos suelen basarse en unidades como pymes u hogares pero, los mismos principios son válidos para la mayoría de otros tipos de datos

de encuestas. Los pasos básicos para decidir si utilizar los datos existentes o recopilar datos nuevos mediante encuestas, se deberá determinar:

- El enfoque del muestreo.
- El tamaño necesario de la muestra.
- Quién recopilará los datos.
- Desarrollo y prueba piloto del instrumento de recopilación de datos.
- Llevar a cabo un trabajo de campo y realizar el control de calidad.
- Procesar y almacenar los datos.

Generalmente, se suele contratar la implementación de estos diversos, asegurando lo esencial que se comprendan su alcance y sus componentes clave para poder gestionar efectivamente una evaluación de impacto de calidad.

Quién recopila los datos

Es indispensable, que la entidad que recopile los datos:

- Siempre se coordine estrechamente con la dirección de introducción de las innovaciones que implemente el programa para garantizar que las operaciones del programa no se pongan en marcha antes de recopilar los datos de línea de base.
- Cuando se necesiten datos de línea de base para el funcionamiento del programa (por ejemplo, datos para un índice de focalización, en el contexto de una evaluación basada en un diseño de regresión discontinua), sea capaz de procesar los mismos rápidamente y transmitirlos a la institución encargada de las operaciones del programa.
- También se requiere una coordinación estrecha de la programación de la recopilación de datos de la encuesta de seguimiento. Por ejemplo, si se ha elegido una implementación con asignación aleatoria, la encuesta de seguimiento debe llevarse a cabo antes de que

el programa se ponga en marcha en el grupo de comparación, para *evitar la contaminación de datos.*

- Debe grantizar la utilización de los mismos procedimientos de recopilación de datos en los grupos de comparación y de tratamiento. A menudo la agencia de implementación tiene contacto solo con el grupo de tratamiento y no está en una adecuada posición para recopilar datos de los grupos de comparación.

Grupos especializados

Utilizar diferentes organismos de recopilación de datos para los grupos de tratamiento y de comparación *es muy riesgoso,* dado que esto es posible que cree diferencias en los resultados medidos, en los dos grupos sencillamente porque los procedimientos de recopilación de datos son diferentes, pero añade ventajas como:

- Garantizar que estos se consideren objetivos, es decir, libres de intereses en juego en los resultados de la evaluación, lo que añade credibilidad al esfuerzo general de evaluación de impacto.
- Garantizar que los encuestados no perciban la encuesta como parte del programa y, de esta manera, se minimiza el riesgo de que los encuestados den respuestas estratégicas intentando aumentar lo que perciben como la posibilidad de participar en un programa.
- Dado que la recopilación de datos comprende una secuencia compleja de operaciones, se recomienda que una entidad especializada y experimentada sea la responsable. Hay pocos organismos ejecutores de programas con suficiente experiencia para recopilar los datos a gran escala y preservar la calidad, ambos criterios necesarios para una evaluación de impacto.
- En la mayoría de los casos, se tendrá que pensar en contratar a una institución local, como una agencia estadística nacional o una empresa o think tank especializados. La contratación de una institución local, como una agencia estadística nacional, puede exponer

a la institución a los estudios de evaluación de impacto y contribuir a mejorar su capacidad, lo cual en sí mismo puede ser un beneficio secundario de la evaluación de impacto.

Posibles inconvenientes

Algunos inconvenientes de contratar un organismos especializado para recopliar los datos, son:

- Las agencias estadísticas nacionales no siempre tendrán la capacidad logística para asumir otros encargos además de sus actividades regulares.
- Es posible que también carezcan de la experiencia necesaria para llevar a cabo encuestas para las evaluaciones de impacto, como la experiencia de efectuar un seguimientoexitoso de los individuos a lo largo del tiempo, o para aplicar instrumentos de encuesta no tradicionales.

Si estas limitaciones aparecen, contratar una empresa independiente o institución especializada en recopilación de datos puede ser lo más práctico.No siempre es necesario que sea la misma entidad la que recopila información de las encuestas de *línea de base* y de *seguimiento*, dado que pueden variar en su alcance. Por ejemplo, en una evaluación de impacto de un programa de capacitación cuya población de interés está compuesta por los individuos que se inscribieron en el curso, la institución a cargo del curso podría recopilar los datos de *línea de base* cuando los individuos se inscriben. Sin embargo, es poco probable que la misma agencia también sea la mejor opción para recopilar información de seguimiento, tanto para los grupos de tratamiento como de comparación. En este contexto, contratar rondas de recopilación de datos por separado tiene sus ventajas, pero se debería hacer un esfuerzo para no perder información entre las rondas, información que será útil para hacer un seguimiento de los hogares o de

los individuos, así como para asegurar que los datos de línea de base y de seguimiento se midan de manera consistente.

Decidir por el grupo especializado más adecuado

Decidir cuál es la mejor institución para recopilar los datos de la evaluación de impacto, implica:

- Sopesar diversos factores como: experiencia en la recopilación de datos, capacidad de coordinar con la agencia ejecutora del programa, independencia, oportunidades para la mejora de capacidades, adaptabilidad al contexto de la evaluación de impacto, etc.
- Costo previsto y,
- La probable calidad de los datos obtenidos en cada caso.

Una manera efectiva de identificar la organización mejor situada para recopilar datos de calidad consiste en:

- Redactar *términos de referencia claros.*
- Solicitar a las organizaciones que presenten propuestas técnicas y financieras.
- La entrega oportuna y la calidad de los datos son cruciales para la fiabilidad de la evaluación de impacto por lo tanto, el contrato para la agencia encargada de la recopilación de datos debe estructurarse con gran cuidado.
- El alcance del trabajo previsto y los productos deben definirse con suma claridad.
- Además, se recomienda introducir *incentivos* en los contratos y vincular esos incentivos a indicadores claros de la calidad de los datos. Por ejemplo, la tasa de falta de respuesta es un indicador clave de la calidad de los datos. Para crear incentivos con el fin de que las agencias de recopilación de datos *minimicen las no respuestas*, el contrato puede estipular un costo unitario para el primer **80%** de la

muestra, un costo unitario superior para las unidades de entre el **80% - 90%** y, una vez más, un costo unitario superior para las unidades de entre el **90% -100%.**

- Como alternativa, redactar un contrato por separado para que la empresa encuestadora realice un seguimiento de los no encuestados.
- Además, el contrato de la empresa de recopilación de datos puede incluir incentivos o condiciones relacionadas con la verificación de la calidad de los datos, como comprobaciones externas o auditorías de calidad de una submuestra de la encuesta de evaluación de impacto.

Los indicadores y su importancia

A lo largo de la cadena de resultados, los indicadores deben estar presentes y medirse, en etapas: intermedias y finales así como las medidas de los beneficios y la calidad de la implementación del programa. Se sugiere el diseño de cuestionarios **(Mejía-Trejo, 2019b).** Así, se recomienda ser selectivo con los indicadores, para:

- Los indicadores a medir para limitar los costos de recopilación de datos, lo que simplifica la tarea de la agencia de recopilación y mejora la calidad de los datos recopilados minimizando las demandas de tiempo para encuestadores y encuestados.
- Evitar la recopilación de indicadores que sean irrelevantes o que probablemente no se utilizará tiene un costo muy alto. Los datos adicionales requieren más tiempo de preparación, capacitación, recopilación y procesamiento.
- Aprovechar la disponibilidad y capacidad de atención limitadas, donde los encuestados, al proporcionar información de calidad cada vez más limitada a medida que la encuesta avanza, los entrevistadores, con los incentivos, tenderán a realizar acciones para ahorrar tiempo, con el fin de cumplir con sus objetivos de la encuesta. Objetivos alineados del programa bien definidos, ayudan a priorizar la información necesaria.

Contar con datos de línea de base es sumamente recomendable, ya que:

- Son esenciales para probar si el diseño de la evaluación de impacto es adecuado, aún cuando se utilice una asignación aleatoria o un diseño de regresión discontinua, donde en principio se pueden usar sencillas diferencias después de la intervención para estimar el impacto de un programa.
- Si la asignación aleatoria no funciona, se pueden utilizar métodos de diferencias en diferencias como alternativa.
- Son útiles como variables de control durante la etapa de análisis, dado que contribuyen a aumentar la potencia estadística y permiten analizar impactos en diferentes subpoblaciones.
- Por último, se utilizan en la mejora del diseño del programa ya que en ocasiones permiten analizar la eficiencia focalizada o proporcionan información adicional sobre los beneficiarios a la agencia que implementa el programa. En algunos casos, la encuesta de seguimiento puede incluir un conjunto más amplio de indicadoresque la encuesta de línea de base.

Cómo hacer que los indicadores midan

El próximo paso consiste en determinar exactamente cómo medir esos indicadores, Una vez que se han definido los datos centrales que se debe recopilar. Dada la vital importancia de esta etapa, donde la medición es reconocida como un arte en sí misma, es preferible que de ella se ocupen los especialistas, entre los cuales se encuentran:

- El equipo de investigación de la evaluación de impacto.
- La agencia contratada para recopilar datos.
- Los expertos de las encuestas.
- Los expertos en la medición de indicadores complejos específicos.

Así, deben ser lo más consistentes posible con las mejores prácticas locales e internacionales, los indicadores de resultado, por lo que se recomienda:

- Investigar cómo los indicadores de interés, han sido medidos en encuestas similares, tanto a nivel local como internacional.
- Utilizar los mismos indicadores (lo que incluye los mismos módulos o preguntas de las encuestas) que garantice la comparabilidad entre los datos preexistentes y los datos recopilados para la evaluación de impacto.
- Si se selecciona un indicador que no se tenga la seguridad que sea plenamente comparable o no esté bien medido, puede limitar la utilidad de los resultados de la evaluación. Es posible, que tenga sentido invertir en recursos adicionales para recopilar el nuevo indicador de resultado que se considere *innovador*, así como una alternativa más establecida.
- Se debe asegurar que todos los indicadores, se midan exactamente de la misma manera para todas las unidades tanto del grupo de *tratamiento* como de *comparación*. Esto es:
    - Diferentes métodos de recopilación de datos tales como la encuesta, entrevista, los e-mail, encuesta telefónica para un grupo y una encuesta presencial para otro,etc. tienen tendencia a crear sesgos, lo cual es un riesgo.
    - Lo mismo sucede con la recopilación de datos en diferentes momentos para los dos grupos, como recopilar datos para el grupo de tratamiento durante la estación de lluvias y para el grupo de comparación durante la estación seca. Por esto, los procedimientos empleados para medir cualquier indicador de resultado, deben formularse con alta precisión. El proceso de recopilación de datos debe ser exactamente el mismo para todas las unidades. En un cuestionario, cada módulo relacionado con el programa debe introducirse sin afectar el flujo o la formulación de las respuestas en otras partes del cuestionario. De hecho, cuando sea posible, es preferible

evitar hacer cualquier distinción entre los grupos de tratamiento y comparación en el proceso de recopilación de datos.

- La agencia recopiladora (o al menos los encuestadores individuales) no debe tener motivos para conocer la condición de tratamiento o de comparación de los individuos en la encuesta.
- Definir la forma de medir los indicadores de resultados, es decir: si se hace mediante encuestas tradicionales basadas en cuestionarios y preguntas auto-reportadas o a través de otros métodos.

En los últimos años, se han producido varios avances para medir resultados o conductas clave que son relevantes en las evaluaciones de impacto. Los avances incluyen el perfeccionamiento de los métodos para recopilar datos auto-reportados mediante cuestionarios, así como técnicas para medir directamente los resultados clave.

Otras fuentes de información

En este caso, se sugieren los que están contenidos en bases de datos institucionales, las que son producidas por asociaciones civiles o de gobierno así como las privadas. Sin embargo, es posible que:

- Se precisen otros datos adicionales que no se encuentran en dichas fuentes para la evaluación de impacto, lo cual depende de la metodología usada. Los datos sobre otros factores que influyan en el resultado de interés es posible que sean necesarios para controlar influencias externas. De hecho, este aspecto es importante cuando se utilizan métodos de evaluación que dependen de más supuestos que los métodos aleatorios.
- En ocasiones, es necesario prever datos sobre los resultados y otros factores a lo largo del tiempo para calcular tendencias, como sucede con el método **DD**. Por ejemplo, si se elige el método **PSM** o **DD**, habrá que recolectar datos sobre una gama muy amplia de caracterís-

ticas para los grupos tanto de *tratamiento* como de *comparación*, lo que hace posible ejecutar una serie de pruebas de robustez.

- Deba desarrollarse para cada evaluación de una matriz que enumere las preguntas de interés, los indicadores de resultado para cada pregunta y la fuente de los datos (Ver **Gráfico 2.1**).
- Deba elaborarse un plan de evaluación de impacto y contar con un preanálisis como base de oportunidades esenciales para definir una lista precisa de indicadores clave requeridos en las evaluaciones de impacto.
- Para aumentar la potencia estadística, en ocasiones es conveniente dar cuenta de otros factores y tendencias anteriores.
- Incluso con la asignación aleatoria, los datos sobre otras características ayudan a estimar los efectos del tratamiento con más precisión, por lo que también pueden ser utilizados para incluir controles adicionales o analizar la heterogeneidad de los efectos del programa en características relevantes.

Sin embargo, se corre el riesgo de que estas no se encuentren actualizadas, o que simplemente, no contengan los indicadores requeridos a la evaluación de impacto, pero que sirven como referentes de diseño de cuestionario requerido.

La recopilación de datos representa el mayor costo de una evaluación de impacto. Para determinar si los datos existentes se pueden utilizar en una determinada evaluación de impacto, se sugiere responder las preguntas de la **Tabla 9.7.**

Tabla 9.7. Preguntas que definen si los datos a emplear son aptos para realizar la evaluación de impacto

El muestreo
¿Se dispone de datos existentes tanto para el grupo de tratamiento como para el grupo de comparación? ¿Las muestras existentes se han extraído de un marco muestral que coincide con la población de interés? Las unidades del marco muestral ¿se han obtenido mediante un procedimiento de muestreo probabilístico?
Tamaño de la muestra
¿Las series de datos son suficientemente grandes para detectar cambios en los indicadores de resultado con suficiente potencia? La respuesta a esta pregunta depende de la elección de los indicadores de resultado, así como de los resultados de los cálculos de potencia.
Disponibilidad de datos de línea de base
¿Los datos existentes disponibles tanto para los grupos de tratamiento como de comparación son anteriores a la implementación del programa o de la innovación que se evalúa? La disponibilidad de datos en línea de base es esencial para documentar el equilibrio en las características previas del programa entre los grupos de tratamiento y de comparación cuando se utilizan métodos aleatorios, y son esenciales para la implementación de diseños cuasi experimentales.
Frecuencia
¿Los datos existentes son recopilados con suficiente frecuencia? ¿Se dispone de ellos para todas las unidades de la muestra a lo largo del tiempo, incluyendo los momentos en que hay que medir los indicadores de resultado según la cadena de resultados y la lógica de la intervención?
Alcance
¿Los datos existentes contienen todos los indicadores necesarios para responder a las preguntas de interés de las políticas, incluyendo los principales indicadores de resultado y los resultados intermedios de interés?
Vínculos con la información de monitoreo del programa
¿Los datos existentes se pueden vincular a datos del monitoreo de la implementación del programa, lo que implica observar qué unidades pertenecen a los grupos de tratamiento y de comparación, y si todas las unidades asignadas al grupo de tratamiento reciben los mismos beneficios?
Identificadores únicos
¿Existen identificadores únicos que vinculen diferentes fuentes de datos?

Fuente: Gertler et al. (2017) con adaptación propia.

Cabe destacar finalmente, que el instrumento de recopilación de datos deba ser probado a través de *pruebas piloto* de manera intensa antes de aplicarlo en campo **(Mejía-Trejo, 2019b).** Realizar pruebas piloto amplias e intensas del instrumento, presenta la oportunidad de su adecuación a los diversos contextos previendo su contenido, formato, opciones de redacción alternativas, así como también los protocolos de recopilación de datos, incluida la tecnología. Permite checar su duración y para verificar que su formato sea suficientemente consistente, congruente, coherente y exhaustivo para producir mediciones precisas de toda la información relevante. Las pruebas de campo constituyen una parte integral de la elaboración de los instrumentos de recopilación de datos y evitar los errores de los mismos **(Mejía-Trejo, 2019a,** ver capítulo 3. Análisis de datos cómo manejar errores).

CAPÍTULO 10.

Lineamientos de Gestión de Proyectos de Evaluación de Impacto

El realizar una evaluación de impacto de un programa de innovación, implica antes que nada, un equipo de investigación capaz de diseñar el programa de innovación aliado a su vez, con un equipo que dirija las políticas de introducción de las innovaciones. Esta naturaleza de los dos equipos demanda una caracterización y comprensión tanto los roles y alcances en responsabilidades. Si bien la evaluación de impacto de un programa de introducción de innovaciones se puede aplicar, y se ha aplicado, en muy diversos contextos, también es posible incluirlo en el marco de un proyecto, esto es, la construcción planeada de productos, servicios de innovadores, así como de infraestructura como aeropuertos, puentes, circunvalaciones, represas, carreteras, minas, oleoductos, puertos, líneas de transmisión, centrales eólicas, así como también proyectos comerciales agrícolas y agroindustriales y la creación de áreas de conservación de la naturaleza. La planificación y la construcción de tales proyectos de innovación tienden a causar impactos sociales. Desde las primeras etapas existe especulación acerca de cómo el proyecto podrá afectar los precios de los bienes muebles, y puede conducir ya sea a un éxodo de gente o, a la inversa, al influjo de personas motivado por el programa de innovación a introducir.

La importancia de los proyectos de innovación en el impacto social

De esta forma, los proyectos pueden crear oportunidades y beneficios para la población, pero al mismo tiempo también pueden tener efectos perjudiciales. En general, nunca son uniformemente buenos o malos, sino que existe una distribución diferencial de costos y beneficios en las comuni-

dades aledañas. Es muy simplista decir que hay ganadores y perdedores, porque las personas se pueden beneficiar y perjudicar al mismo tiempo. Es necesario llevar a cabo una adecuada gestión que garantice que los beneficios de los proyectos se maximicen y que se eviten, o minimicen, los impactos negativos de manera constante durante el ciclo del proyecto. La evaluación de impacto social es un proceso de introducción de innovaciones puede contribuir enormemente a garantizar el logro de beneficios y la prevención del daño.

En este sentido, será sugerible elaborar a nivel nacional, el Plan de Gestión de Impacto Social (**PGIS**) consistente en integrar la gestión de los impactos sociales a lo largo del proceso de identificación de oportunidades, contratación y ejecución de los proyectos de innovación, con un enfoque proactivo de compromiso con las comunidades donde opera. Lo anterior, con el fin de monitorear, rendir informes, evaluar, revisar y responder proactivamente al cambio. La gestión adaptativa es parte importante de la gestión de impactos sociales. Los gobiernos e inversores en proyectos están requiriendo cada vez más estos planes de gestión de impacto social. En general se deben elabora al momento de preparar las declaraciones de impacto para obtener las aprobaciones del proyecto y luego se los actualiza periódicamente. Idealmente, debería haber correspondencia con los sistemas de gestión interna de las empresas o, en su defecto, deberían ser alimentados; alternativamente, las empresas podrían incluir un sistema de gestión de impacto social en su planificación general **(AIEI, 2015).**

Los principios internacionales de la evaluación del impacto social consideran que los impactos sociales incluyen todos los aspectos asociados con una intervención planeada (esto es, un proyecto) que afectan o involucran a laspersonas, ya sea directa o indirectamente. Específicamente, un impacto social es algo que se experimenta o se siente, en el sentido perceptual (cognitivo) o corporal (físico) a todos los niveles, por ejemplo, a nivel de la persona como individuo, de unidad económica (familia/hogar), de grupo social (círculo de amigos), de lugar de trabajo (una empresa o entidad de gobierno), o más generalmente de comunidad/sociedad. Estos diferentes

niveles se ven afectados de diferentes maneras por un impacto o por una acción que causa impacto **(AIEI, 2015).**

Existen conceptos que son importantes diferenciar y precisar en los proyectos de introducción de innovaciones con impacto social, tales como:

- El objetivo de todos los proyectos debe ser el desarrollo social sostenible.
- Un concepto importante es la distinción entre un proceso de cambio social y un impacto social. No todos los procesos que inducen cambios en una comunidad necesariamente causan impactos sociales.
- Si bien se describe a veces la evaluación de impacto social como una forma social de la evaluación de impacto ambiental, existen muchas diferencias entre ambas. A saber: los impactos ambientales tienden a ocurrir a partir de que se mueve la primera piedra, mientras que los impactos sociales pueden ocurrir en cuanto comienza a correr el rumor de que algo puede llegar a ocurrir.
- La evaluación de impacto social ayuda a establecer una base de empleo y de aprovisionamiento local.
- La evaluación de impacto social no es lo mismo que participación pública.
- Un rumor, ya sea fundado o no, también puede circular y amplificar los temores y ansiedades de la población, más allá de que el proyecto se concrete o no. Temor y ansiedad, como todos los impactos percibidos, son impactos sociales reales que la población experimenta y no se los debe descartar, sino por el contrario deben gestionarse de manera efectiva. Los impactos sociales rara vez son relaciones únicas de causa y efecto, sino que son complejos patrones de vías de impacto que se intersectan.
- Los impactos sociales también pueden ser causados por proyectos que crean falsas expectativas en una comunidad, ya sea que el personal del proyecto promete intencionalmente y de manera inapropiada cosas que no ocurrirán, o inadvertidamente mediante una gestión

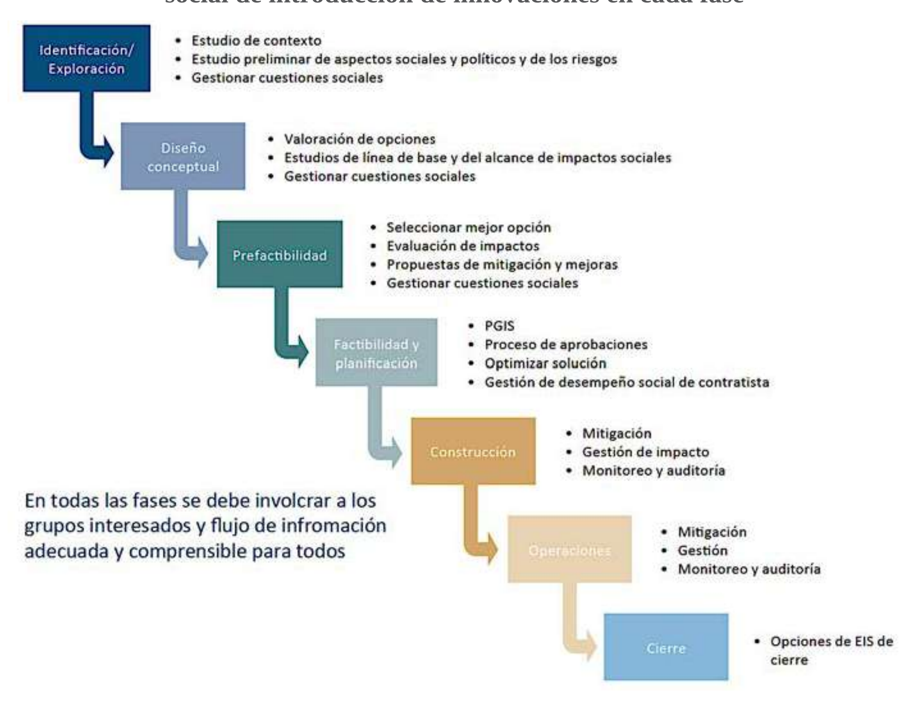
pobre de las expectativas y permitiendo que los rumores aumenten las expectativas.
- Los proyectos invariablemente aumentan la desigualdad en una comunidad, razón por la cual es realmente importante que los proyectos contribuyan con iniciativas de inversión social local.
- Elaborar una evaluación de impacto social es buen negocio y es bueno para el negocio.
- Se deben considerar los derechos humanos.
- Pueblos indígenas, tradicionales, tribales y otros vinculados a las tierras deben ser reconocidos y recibir particular atención.
- El desplazamiento y el reasentamiento son la principal causa de perjuicio para las comunidades afectadas y constituyen uno de los mayores riesgos para los proyectos.
- El contenido local permite crear valor compartido.
- Se debe planificar el cierre al inicio del poryecto.
- Se debe ejercer ética en la evaluación de impacto social par al aintroducción de innovaciones.
- Dada la complejidad, es práctico crear mapas mentales en el proceso de la determinación del alcance.

Sociedad participativa en las etapas de proyectos de innovación

Un cambio importante que ha ocurrido en la entrada del siglo XXI, es la mayor concientización de que abordar los impactos sociales requiere la gestión activa de los aspectos sociales desde el primer momento de un proyecto, mucho antes de que sea necesaria la aprobación reglamentaria. Esto es, las comunidades no quierne que se les impongan los impactos de los proyectos quieren participar activamente en el desarrollo conjunto y quieren obtener beneficios de los proyectos del sector privado. El valor compartido entonces es una consideración importante. También se ha caído en la cuenta que las entidades regulatorias y de monitoreo hoy ya no son únicamente

entidades gubernamentales, sino que incluyen a las ONG locales y globales (en particular, las ONG de vigilancia), asociaciones industriales internacionales, los sectores de finanzas y seguros, y las propias comunidades afectadas por proyectos. De esta manera, la evaluación de impacto oscila de un proyecto de introducción de innovaciones involucra a todos los grupos interesados donde si bien el ciclo de un proyecto a menudo se representa como un proceso lineal, la realidad no es tan directa: los proyectos no necesariamente transitan sobre ruedas de una fase a otra, y pueden detenerse en cierta fase o ir hacia atrás a una fase anterior. La evaluación de impacto social de introducción de proyectos de innovación es el proceso de gestión de las cuestiones sociales de los proyectos.Ver **Figura 10.1.**

Figura 10.1. Ciclo de proyecto y el rol potencial de la evaluación de impacto social de introducción de innovaciones en cada fase



Fuente: AIEI (2015).

La evaluación de impacto social en la introducción de innovaciones ha sido mucho más efectiva en influenciar el diseño de un proyecto y mejorar los resultados para las comunidades cuando sus proponentes la han considerado parte clave de sus procesos de planificación y gestión del riesgo. Aun cuando la mayor parte de los consultores defienden a las comunidades, existe un evidente desequilibrio de poder entre el proponente y la comunidad. Un modelo que se puede aplicar en contextos de baja confianza entre el proponente y la comunidad, consiste en que el *proponente brinde fondos a las comunidades receptoras para que puedan contratar a sus propios* consultores en la evaluación de impacto. De esa manera, esa evaluación de impacto social de introducción de innovaciones *basada en la comunidad* permitiría que la comunidad tenga su propia información al respecto. Se puede decir que este modelo apoyaría un proceso de negociación que fuera consistente con el espíritu del consentimiento libre, previo e informado. **Ver Figura 10.2.**

Figura 10.2. Fases de la evaluación del impacto social

Fuente: AIEI (2015).

Las tareas se presentan en orden cronológico aproximado, pero cabe mencionar que cada una sirve de base para las demás, y a medida que la información se acumula en la EIS y que se obtiene nueva información, tanto las decisiones tomadas en una fase más temprana del proceso sobre el alcance, el área o la influencia, como los actores, deberán ser evaluados nuevamente. El proceso es, por lo tanto, iterativo. **Ver Tabla 10.1.**

Tabla 10.1. Las 26 tareas que comprenden la evaluación de impacto social en la introducción de innovaciones

Fase 1: Entender los problemas
1. Obtener una correcta comprensión del proyecto propuesto, incluidas todas las actividades complementarias necesarias para apoyar el desarrollo y el funcionamiento del proyecto.
2. Aclarar las responsabilidades y roles de todos quienes participan o están vinculados a la EIS, incluidas las relaciones con los demás estudios especializados que se han emprendido, y determinar las leyes nacionales o normas internacionales que deben cumplirse.
3. Identificar el "área de influencia social" preliminar del proyecto, las comunidades que se verán probablemente afectadas y beneficiadas (cercanas y distantes) y las partes interesadas.
4. Lograr una buena compresión de las comunidades que podrían verse afectadas por el proyecto, preparando un perfil o descripción de la comunidad que incluya: (a) un análisis minucioso de los actores interesados; (b) un examen del marco sociopolítico; (c) una evaluación de las diferentes necesidades, intereses, valores y aspiraciones de los diversos subgrupos de las comunidades afectadas, incluido un análisis de género; (d) una evaluación de su historial de impactos, es decir su experiencia con proyectos anteriores y otros acontecimientos históricos; (e) un análisis de las tendencias que se dan en esas comunidades; (f) un examen de los recursos, fortalezas y debilidades de las comunidades; y (g) como opción, los resultados de una encuesta de opinión. Por lo general, esta tarea se denomina elaboración de perfil.
5. Informar plenamente a los miembros de la comunidad acerca de: (a) el proyecto; (b) proyectos similares en otros lugares, de modo de darles una idea de cómo podrían verse afectados; (c) cómo pueden participar en la evaluación de impacto social; (d) sus derechos en materia de procedimientos dentro del marco regulatorio y de desempeño social del proyecto; y (e) su acceso a los mecanismos de reclamos y de intercambio de información.
6. Implementar procesos participativos y espacios de deliberación inclusivos de manera tal de ayudar a los miembros de la comunidad a: (a) entender cómo se verán afectados por el impacto; (b) determinar la aceptabilidad de posibles impactos y beneficios propuestos; (c) tomar decisiones informadas acerca del proyecto; (d) facilitar la visión de la comunidad acerca del futuro deseado; (d) contribuir a los planes de mitigación y de monitoreo; y (f) prepararse para el cambio.
7. Identificar las cuestiones sociales y de derechos humanos que podrían ser una potencial fuente de preocupación (alcance o scoping).
8. Recopilar datos pertinentes para elaborar la línea base para los principales problemas sociales.

Fase 2: Predecir, analizar y evaluar las vías de impacto posibles
9. Por medio del análisis, determinar los cambios e impactos sociales que probablemente ocasione el proyecto y sus diversas alternativas. **10.** Considerar minuciosamente los impactos indirectos. **11.** Examinar de qué manera el proyecto contribuirá a los impactos acumulativos que padecen las comunidades afectadas. **12.** Determinar como los grupos afectados podrian responder a los impactos . **13.** Establecer la importancia de los cambios pronosticados (es decir priorizarlos). **14.** Contribuir activamente al diseño y a la evaluación de alternativas de proyecto, incluso la de no continuar.
Fase 3: Desarrollar e implementar estrategias
15. Identificar maneras de tratar los impactos negativos potenciales (utilizando la jerarquía de mitigación). **16.** Desarrollar e implementar maneras de maximizar los beneficios y las oportunidades relacionadas con el proyecto. **17.** Desarrollar estrategias para apoyar a las comunidades a hacer frente al cambio. **18.** Desarrollar e implementar mecanismos adecuados de reclamos y de intercambio de información. **19.** Facilitar un proceso de creación de acuerdos entre las comunidades y el promotor del proyecto que lleve a la redacción de un Acuerdo de Impactos y Beneficios (**AIB**). **20.** Ayudar al proponente a facilitar el aporte de información de las partes interesadas y a redactarun Plan de Gestión de Impacto Social (**PGIS**) que ponga en práctica los beneficios, medidas de mitigación, sistemas de monitoreo y mecanismos de gobernanza que hayan sido convenidos en el Acuerdo de Impactos y Beneficios (**AIB**), como asimismo planes para tratar aquellas cuestionesno previstas que pudieran surgir. **21.** Implementar procesos que permitan a los proponentes, a las autoridades gubernamentales y a las partes interesadas de la sociedad civil poner en práctica los mecanismos incorporados en el Plan de Gestión de Impacto Social (**PGIS**) y los Acuerdos de Impactos y Beneficios (**AIB**), y desarrollar e integrar sus respectivos planes de acción de gestión en sus propias organizaciones, establecer los respectivos roles y responsabilidades mediante la implementación de esos planes de acción, y mantener un rol permanente en el monitoreo. **22.** Ayudar al proponente a desarrollar e implementar planes de desempeño social permanentes que traten de las obligaciones de los contratistas incorporadas en el **PGIS**.
Fase 4: Diseñar e implementar programas de monitoreo
23. Desarrollar indicadores para efectuar un monitoreo de los cambios en el tiempo. **24.** Desarrollar un plan de monitoreo participativo. **25.** Examinar de qué manera se implementará la gestión adaptativa y analizar la implementación de un sistema de gestión social. **26.** Realizar una evaluación y una revisión periódica (auditoría).

Fuente: AIEI (2015).

La empresa y su compromiso social

A pesar de que en ciertos lugares pueda ser un requisito legal, la razón moral o normativa consiste en que las comunidades esperan de las empresas una evaluación de impacto social adecuada y por lo tanto integra sus obligaciones en materia de responsabilidad social: elaborar una evaluación de impacto social es lo que se debe hacer. No obstante, si bien este argumento tiene cierta justificación altruista, no necesariamente convence a aquellas empresas que creen que su obligación es maximizar retornos para sus accionistas manteniéndose dentro de lo permitido por la ley. De manera similar, el argumento basado en el propio interés de elaborar una evaluación de impacto social a la introducción de innovaciones adecuadamente para que las empresas consigan su licencia social para operar tampoco convence del todo a los ejecutivos de las empresas, porque según esta perspectiva la EIS es vista solamente como un costo/gasto necesario para obtener la aprobación y el permiso social. Según esta perspectiva, siempre se intentará reducir los gastos en especial porque no hay una conexión directa percibida o fácilmente confirmada entre el costo de la evaluación y su retorno en forma de aprobación social. Por el contrario, los defensores de la evaluación de impacto social sostienen que elaborarlo apropiadamente no se debería considerar un costo sino una inversión en gestión de riesgo. Disminuye los posibles gastos a futuro de una empresa ya que identifica posibles problemas y por tanto reduce los costos por posibles litigios, retrasos en aprobaciones, control de medidas de protestas o por manejar situaciones de violencia contra el personal y/o los bienes, así como las pérdidas comerciales por perjuicio de la reputación. Reducir el riesgo también lleva a reducir costos de capital y por consiguiente a aumentar el valor para los accionistas. De este modo, existen muchas razones por las que elaborar una evaluación de impacto social en la introducción de innopvaciones de forma apropiada tiene sentido para las empresas, ya que agrega valor a todos los impulsores de empresas**. Ver Figura 10.3.**

Ver Figura 10.3. Impulsores de valor empresarial a los que la evaluación de impacto social en la introducción de innovaciones agrega valor

Fuente: AIEI (2015).

Para operar en un país, las empresas, en particular grandes empresas extranjeras multinacionales, obtienen un permiso legal específico, pero también necesitan procurar y mantener la licencia social para operar. Para obtenerla, necesitan contribuir positivamente con el país y, lo que es más importante, con las comunidades locales en las que operan, más allá de cualquier tributo o regalía que se les pueda requerir pagar. Los derechos a acceder a oportunidades lucrativas provienen de las oportunidades que las empresas brindan y de las inversiones para contribuir al desarrollo social en el país receptor y en sus comunidades locales. A cambio, se las considera empresas confiables, socialmente responsables y elegidas para realizar ese proyecto, lo cual es positivo para su reputación y para acceder a mercados en otros lugares: sin duda una buena parte de la propuesta de valor compartido.

Para obtener y mantener una licencia social genuina se requiere que el proyecto reflexione sobre su contribución al desarrollo social. El desarrollo social significa más que solamente proporcionar unos pocos empleos y brindarfondos para una nueva escuela o piscina pública: se requiere que el proyecto se asocie con las comunidades locales para ser una fuerza de cambio social positivo y desarrollo social beneficioso. Debe ser un proceso participativo de cambios sociales diseñado para mejorar el bienestar de la comunidad en su conjunto y especialmente de los grupos vulnerables, desfavorecidos o marginados dentro de una región.

Más que tratarse de beneficios para individuos *per se*, el desarrollo social trata de facilitar el cambio en instituciones y en la sociedad para reducir la exclusión social y la fragmentación, promover la inclusión social en la introducción de innovaciones y la democratización, y para fortalecer la capacidad institucional y de gobernanza. El desarrollo social ve más allá de los problemas y las deficiencias para centrarse en mejorar las capacidades de las personas y las instituciones y en cómo se las puede fortalecer. No obstante, las empresas deben ser claras y enfocadas en sus intervenciones para garantizar que no absorben o usurpan el rol y las responsabilidades de las autoridades locales.

La inversión social consiste en las contribuciones financieras y en especie que un proyecto hace a la comunidad local para el desarrollo social. Lo ideal sería que esos recursos no se desperdiciaran en actividades puntuales y no sostenibles, sino que contribuyeran de manera concreta al desarrollo social y económico de mediano y largo plazo. Para esto, debe haber un proceso de evaluación que seleccione y priorice las opciones de inversión social. El término inversión social estratégica se emplea cuando, además de apoyar los resultados en materia de desarrollo social para la comunidad local, existe un claro argumento para que la empresa ponga los fondos a disposición. Contribuir con los programas de capacitación que fortalecen las capacidades de la mano de obra local para que proporcionen trabajo o servicios para el proyecto es un claro ejemplo de valor compartido.

Los objetivos de desarrollo sostenible de la ONU y el capital comunitario

Los objetivos de desarrollo social apropiados variarán según el contexto particular de aplicación. Identificar esos objetivos debe ser un proceso participativo guiado por la comunidad. En términos generales, es pertinente tomar en cuenta los 17 Objetivos de Desarrollo Sostenible de las Naciones Unidas **(ONU CEPAL, 2015). Ver Imagen 10.1.**

Identificar y evaluar opciones posibles para la inversión social también requiere pensar sobre las fortalezas y debilidades dentro de una comunidad. Un marco de referencia empleado comúnmente es el enfoque de *capitales comunitarios* (a veces conocido como el *modelo de la pirámide*) que apuntala el enfoque de medios de subsistencia sostenibles y sus variantes. El enfoque de los medios de subsistencia sostenibles considera las capacidades, los recursos para conseguir sustento (activos, capitales) y las estrategias de sustento (actividades) que las personas emprenden para ganarse la vida y vivir según eligen hacerlo.

Imagen 10.1. Los 17 Objetivos de Desarrollo Sostenible de la ONU



Fuente: ONU CEPAL (2015).

En el centro del modelo se encuentra la noción de que todos los recursos o activos comunitarios se pueden representar como serie de capitales. La evaluación de las estrategias de inversión social puede tener en cuenta esos capitales y cómo fortalecer uno o más de aquellos aumenta el bienestar general en la comunidad. **Ver Tabla 10.2.**

Tabla 10.2. Capital comunitario

• **Capital natural:** comprende flujos y stocks de activos provenientes de recursos naturales (esto es, servicios de ecosistemas) como recursos alimentarios y agrícolas, recursos forestales, reservas minerales, suelo, agua,humedales y reservas pesqueras. • **Capital físico** (también llamado producido, manufacturado o construido): comprende la reserva de equipamientos, planta física (esto es, fábricas), infraestructura (esto es, carreteras, aeropuertos, hospitales, escuelas) y otros recursos productivos propiedad de individuos, del sector comercial o del país, así como también los sistemas de gestión necesarios para que funcionen. • **Capital financiero:** los recursos financieros disponibles para las personas, como sus ahorros y el acceso a crédito. También da cuenta de toda deuda o hipoteca que se posea. • **Capital humano:** incluye los niveles de conocimiento y habilidades, educación formal, salud y nutrición de los individuos, así como su motivación y aptitud. • **Capital social**: a veces definido simplemente como redes sociales y confianza solamente, incluye también las reglas sociales, normas, obligaciones y convenios de reciprocidad incorporados en las relaciones sociales, estructuras sociales y en los arreglos institucionales de la sociedad. • **Capital político o institucional:** hace referencia a la existencia y funcionamiento efectivo (esto es, capacidad) de los mecanismos de gobernanza de la sociedad, a las instituciones de gobernanza en sí mismas, y a las normas, reglas, reglamentos que aplican y que imponen. • **Capital cultural y espiritual:** incluye la manera en que las personas conocen el mundo así como su lugar en este mundo, y también cómo actúan en él. También refiere a la medida en que la cultura local, las tradiciones, el idioma, etc., promueven o entorpecen el bienestar, la inclusión social y el desarrollo social. El capital espiritual contribuye a mantener un balance a través de los diferentes capitales y a permanecer en contacto con valore muy arraigados y con aquello que da sentido a la vida. El capital cultural incide en qué voces se escuchan y a cuáles se atiende, qué voces influyen en qué áreas, y cómo la creatividad, innovación e influencia surgen y se nutren.

Fuente: AIEI (2015).

Existen múltiples formas de capital (activos, recursos) y muy diferentes maneras de agruparlas y definirlas. Lo que se incluya debe depender del contexto de aplicación. *Los capitales individuales se consideran metáforas* y se emplean genéricamente en vez de ser definidos estrictamente e interpretados de forma restrictiva. El concepto de capitales se puede aplicar en diferentes niveles de análisis: para aplicarlo a un individuo, un hogar, una

comunidad local, o una región. El enfoque de capitales se creó originalmente para comprender las estrategias de medios de subsistencia de los individuos que vivían en comunidades rurales pobres en países en desarrollo. Hoy se aplica a una gran variedad de situaciones.

Componentes del equipo de gestión de la evaluación de impacto

Básicamente se fundamenta en los equipos de investigación y equipo de políticas en la introducción de innovaciones, como se muestra en la **Tabla 10.3.**

Tabla 10.3. Equipos gestión de la evaluación de impacto

Equipo de investigación	
Función principal	**Actividades principales**
Es responsable de la calidad técnica y la integridad científica del trabajo de evaluación	• Responsabilidades abarcan el diseño de la investigación, la calidad de los datos y el análisis. Se integra de diferentes personas.
Investigador principal (coordinador de proyecto)	
Función principal	**Actividades principales**
Trabaja con los responsables de las políticas y los encargados del programa para: establecer objetivos clave, cuestiones de políticas, indicadores y necesidades de información de la evaluación (uso de una teoría del cambio, como una cadena de resultados	• Decide la metodología de evaluación de impacto. • Desarrollar el plan de evaluación. • Conforma el equipo de investigación. • Registra la evaluación de impacto. • Obtiene aprobaciones de la junta de revisión institucional. • Prepara plan de evaluación, y preanálisis. • Dirige el análisis de los resultados. • Colabora con el equipo de políticas de introducción de innovaciones para divulgar los resultados. • Trabaja efectivamente con todo el equipo de evaluación así como la organización que recopila los datos. • Diversos investigadores pueden trabajar con el investigador principal o como co-investigadores principales para liderar o apoyar trabajos analíticos específicos de los elementos, como el muestreo, las evaluaciones cualitativas o el análisis de costo-efectividad.

Gestor de la evaluación o coordinador del trabajo de campo	
Función principal	**Actividades principales**
Trabaja directamente con el investigador principal en la implementación diaria de la evaluación.	• Trabajar con los encargados del programa y los responsables de las políticas de introducción de innovaciones de dicho equipo. • Supervisar el trabajo de campo cuando se recopilan los datos primarios. • Es el enlace con el investigador principal cuando este carece de una base local.
Experto en estadística y muestreo	
Función principal	**Actividades principales**
Propone y orienta el trabajo de cálculo de potencia, muestreo y estadística a emplear.	• Llevar a cabo cálculos de potencia. • Seleccionar la muestra, comparar los resultados de la muestra real con los de la muestra diseñada. • Ofrecer recomendaciones sobre implicaciones para el análisis en coincidencia con el plan de preanálisis.
Un equipo de recopilación de datos	
Función principal	**Actividades principales**
Elaborar los instrumentos de recopilación de datos y los manuales y libros de código correspondientes.	• Recopilar, digitar y limpiar los datos. • Entregar una base de datos *limpia* y documentada cuando se requiera una recopilación de datos primarios.
Equipo de políticas de introducción de la innovación	
Función principal	**Actividades principales**
Establecen y dirigen la agenda de investigación, definen la pregunta fundamental que aborda el estudio, aseguran los recursos adecuados para el trabajo, y aplican los resultados a las políticas.	• Articular con claridad los objetivos tanto del programa como de la evaluación, así como la teoría del cambio y los principales indicadores de interés. • Supervisar la dimensión del tamaño del efecto mínimo relevante para las políticas de los indicadores de resultado de interés (ver capítulo 2). • Tener conocimiento del diálogo de políticas y de los contactos con las principales partes interesadas. • Asegurar que la evaluación se diseñe para ser lo más relevante posible para las políticas de introducción de innovaciones. • Garantizar que los interesados y los responsables de la toma de decisiones adecuados participen en momentos clave del proceso de evaluación. • Estar en contacto permanente con el equipo de investigación.

Asistentes del programa	
Función principal	**Actividades principales**
Trabajan estrechamente con el equipo de investigación para alinear el diseño de evaluación con la implementación del programa.	• Verificar que el diseño de evaluación se base en información precisa de la operación del programa. • Implementar el programa según lo planificado. • Gestionar el presupuesto de evaluación. • Apoyar al equipo de investigación a supervisar el trabajo de campo en la recopilación de datos.

Fuente: Gertler et al. (2017) con adaptación propia.

Una alianza efectiva entre ambos equipos es vital para asegurar la credibilidad técnica y el impacto de una evaluación de programa de introducción de innovaciones. Una buena práctica para incentivar la colaboración, consiste en asegurar que los equipos locales participen activamente en la gama más amplia de actividades de evaluación. Tales actividades puede ser como talleres conjuntos, publicaciones conjuntas que den garantía de la capacitación y el desarrollo de capacidades, consiguiendo investigadores locales bien situados para que contribuyan de manera adecuada y sirvan como un canal válido entre los equipos de investigación y de políticas de introducción de innovaciones.

De los investigadores, también es posible elevar su compromiso con el conocimiento y uso de una metodología de evaluación robusta y defendible, y tendrán que asegurar su participación en el diseño de la evaluación de impacto, en el análisis de los datos y en generar investigación primaria que cumpla con las normas científicas vigentes en las publicaciones académicas. El reto es unificar los enfoques inter, trans y/o multidisciplinarios que se manejenen la introducción de innovaciones que impliquen la relación de áreas muy diversas como la educación, la ingeniería, la salud, la economía o la sociedad, por mencionar sólo un caso. Los diferentes intereses del equipo de políticas de introducción de innovaciones y del equipo de investigación pueden crear tensiones que hay que entender y gestionar. Así, se puede concluir ciertos principios rectores a tomar en cuenta para coordinar a los equipos, tales como:

- Hacerlos participar desde el comienzo, para maximizar las opciones del diseño de la evaluación y asegurar una asociación efectiva entre los equipos.
- Tener claro un plan de evaluación de impacto desde el comienzo, así como las reglas de operación.
- Comprender los roles, responsabilidades y motivaciones de las diferentes partes interesadas y darles la oportunidad de participar en la evaluación.
- Participar a lo largo de la evaluación para asegurar una alineación adecuada entre la evaluación y la intervención que se evalúa.
- Reconocer y gestionar los riesgos y beneficios, dejando claro qué pueden y no pueden hacer las evaluaciones de impacto.
- Valorar la transparencia y asegurar la objetividad; estar preparados para respetar los resultados, sean buenos o malos.

Así, se propone una descripción general del plan de evaluación, como el que se muestra en la **Tabla 10.4.**

Tabla 10.4. Descripción general de un plan de evaluación de impacto

1. **Portada.**
2. **Contraportada.** Descripción de la publicación que indica los autores (es decir, los nombres de los responsables de hacer el trabajo y redactar el informe), el editor, la fecha de publicación y otros datos que permiten establecer la índole y el objeto del documento.
3. **Resumen ejecutivo.** Breve reseña de temas clave y hallazgos.
4. **Informe de revisión por expertos.** Una carta/informe de cualquier experto o par revisor (o quizás una declaración conjunta si hubo varios revisores) que indica cómo se realizó la revisión, qué limitaciones se aplicaron a los revisores y todo tipo de comentarios, inquietudes y recomendaciones de los revisores. Podría ser apropiada también una respuesta de los autores a la revisión.
5. **Introducción.** Una introducción general al informe que deje en claro su objetivo, que podría incluir una breve descripción general acerca de la manera en que el documento se vincula a la bibliografía/filosofía.
6. **Resumen del proyecto.** Una correcta descripción del proyecto y todas las actividades complementarias, de modo que los lectores puedan hacerse una idea de aquel. Si existen alternativas u opciones del proyecto, podrían explicarse.
7. **Metodología (descripción de la intervención).** Una descripción del diseño general de la EIS, de los métodos empleados, de qué procesos de compromiso comunitario se emplearon, y de qué manera se tomaron en consideración y se abordaron las cuestiones éticas. Quizás definiciones y/o un análisis de los principales conceptos. Se esperaría que se incluyera aquí algún vínculo a la evaluación de impacto social de introducción de innovaciones y la bibliografía sobre investigación social. Debería incorporarse un análisis de los mecanismos de gobernanza para la realización de la evaluación de impacto social de introducción de innovaciones. Es importante que comprenda también las limitaciones de la metodología aplicada, incluidas las decisiones para reducir o expandir el alcance durante el transcurso de la evaluación de impacto social de introducción de innovaciones.
8. **Marco legal y normas aplicables.** Un análisis del marco o marcos legales y de la legislación, reglamentos y directrices vigentes que se aplican a cada caso en particular. Esto incluiría no sólo la legislación/reglamentos locales, las respectivas instituciones y su responsabilidad con respecto al proyecto, sino también una mención de las normas internacionales tales como las Normas de Desempeño de la Corporación Financiera internacional, las directivas de las organizaciones industriales nacionales, y una referencia a la presente guía.
9. **Perfil o descripción de la comunidad y línea de base social.** Si se va a incluir como apéndice un perfil comunitario extendido y una línea base social, deberá incluirse aquí al menos un resumen de las características y los principales grupos interesados; como alternativa, podrá incluirse aquí el perfil comunitario y la línea base social. También deberán analizarse las principales cuestiones históricas. Además, deberán incluirse los aspectos principales del medio ambiente físico que pudieran tener que ver con la compresión del contexto.

10. **Objetivos de la evaluación**
    a. Hipótesis, teoría del cambio, cadena de resultados
    b. Preguntas de políticas
    c. Indicadores de resultados clave
    d. Riesgos
11. Diseño de evaluación
12. Muestreo y datos
    a. Estrategia de muestreo
    b. Cálculos de potencia
13. **Visión general del plan de preanálisis o Informe de la determinación del alcance** . Un informe sobre todos los potenciales impactos considerados en la fase d evaluación. Debe aclararse la disposición de cada impacto que se considere Cuando se presenta como informe separado, deberá incluirse un resumen. Como alternativa, esto puede ser un apéndice.
14. **Plan de recopilación de datos de datos piloto así como de campo/Análisis**
    a. Encuesta de línea de base
    b. Encuesta(s) de seguimiento
15. Implementación del plan de evaluación y productos por entregar
    a. Informe de línea de base
    b. Informe de evaluación de impacto
    c. Nota informativa sobre políticas
    d. Bases de datos, diseño y protocolos de análisis plenamente documentados
16. **Listado priorizado de los principales impactos sociales.** Se trata de un listado de los impactos residuales, con un análisis de la manera en que se ven afectados los diferentes actores interesados. Debería ponerse un particular énfasis en los pueblos indígenas, mujeres y grupos vulnerables.
17. **Reasentamiento (resumen)**. Si se requiere reasentamiento, o si se producirá desplazamiento físico o económico, una breve descripción de la manera en que se emprenderá el proceso de reasentamiento, qué indemnización se brindará y cómo se determinará, y qué medidas se adoptarán para la recuperación y mejoramiento de los medios de subsistencia. Se requerirá la presentación de un plan de acción de reasentamiento totalmente desarrollado, en forma de documento separado.
18. **Resumen de las dedidas de mitigación y gestión.** Deberá proveerse una lista de las medidas de mitigación y otras medidas de gestión para abordar las cuestiones sociales. Debería haber una estimación de costos y un plazo para la implementación de las medidas de mitigación propuestas.
19. **Plan de monitoreo y plan de contingencia (gestión adaptativa**). Debe presentarse un plan sobre la manera en que se emprenderá el monitoreo: lo que se monitoreará, cómo se monitoreará, con qué frecuencia y quién es el responsable, como asimismo de qué manera responderá la empresa si se superara un umbral admisible.
20. **Descripción de los beneficios.** Se trata de un informe acerca de los probables beneficios del proyecto para las comunidades locales, incluidas todas las acciones de inversión social propuestas, el contenido local y las estrategias de contratación a nivel local.

21. **Estrategia de compromiso comunitario permanente y mecanismos de reclamo.** Una descripción de los procesos previstos de compromiso comunitario permanente. Además, una descripción de los mecanismos de reclamos que se introducirán y qué procesos se utilizarán para la gestión de las quejas.
22. **Mecanismos de gobernanza.** Un análisis de los mecanismos de gobernanza que se aplicarán a los procesos de compromiso comunitario permanente, los mecanismos de reclamos, el proceso de monitoreo, y para garantizar la aceptabilidad permanente del programa de investigación social.
23. **Plan de divulgación**
24. **Protocolos éticos sobre protección de sujetos humanos**
    a. Asegurar el consentimiento informad
    b. Obtener aprobación de la junta de revisión institucional
25. **Calendario**
26. **Presupuesto y financiamiento**
27. **Composición y roles del equipo de evaluación**
28. **Referencias**. Una lista de todas las referencias que se utilizan en el informe, y todas las principales referencias que informaron acerca del diseño de la investigación para la EIS.
29. **Apéndices**. Los apéndices variarán de un proyecto al otro y se verán afectados por lo que esté comprendido en el cuerpo del informe, pero podrán incluir: cuestionarios, cronogramas de entrevistas, plantillas de formularios de consentimiento, un perfil comunitario extendido, datos de la línea base, un informe sobre la determinación del alcance (es decir, un listado de todas las cuestiones que se consideren posibles impactos sociales).

Fuente: Gertler et al. (2017) con adaptación propia.

Los equipos deberán considerar la gestión de los aspectos sociales de los proyectos y por lo tanto toda la extensión del trabajo de los profesionales que elaboran la evaluación de impacto social de proyectos de innovación incluye potencialmente las actividades enumeradas abajo. Nótese, sin embargo, que estas tareas no se realizan simultáneamente y que lo que el profesional debería hacer en cada caso específico dependerá del contexto particular. **Ver Tabla 10.5.**

Tabla 10.5. Actividades de ambos equipos (investigación y políticas de introducción de innovaciones) a supervisar

- Elaborar una evaluación ex-ante de los posibles impactos sociales.
- Asegurar que los aspectos sociales de los impactos ambientales y de biodiversidad de los cambios sociales se incluyan en los planes de gestión y evaluación de impacto.
- Compilar un perfil o descripción de la comunidad (esto es, describir el contexto social local).
- Desarrollar una línea base social pertinente para la toma de decisiones y la documentación de los cambios sociales (esto es, recolectar datos de variables sociales clave para documentar el estado pre impacto).
- Identificar e implementar cambios al proyecto y tomar otras medidas para mitigar impactos sociales.
- Planificar el reasentamiento y/o la indemnización para quienes vayan a ser desplazados en instancias en las que no haya otra opción viable.
- Preparar un marco de políticas de reasentamiento, un plan de acción de reasentamiento, un plan de restablecimiento y mejora de medios de subsistencia, y garantizar que estos se integren en los planes de desarrollo del proyecto y en sus cronogramas.
- Elaborar evaluaciones de debida diligencia en materia de derechos humanos y de impacto sobre los derechos humanos, con la participación de expertos en derechos humanos en caso de ser necesario.
- Identificar maneras de mejorar los beneficios del proyecto.
- Identificar a los actores clave, sus intereses, relaciones e impactos potenciales en relación al proyecto.
- Facilitar los procesos genuinos de compromiso comunitario consistentes con el espíritu de consentimiento libre, previo e informado.
- Cuando exista participación de comunidades indígenas, asistir en el proceso de cumplir los requisitos formales del consentimiento libre, previo e informado.
- Asistir a las comunidades afectadas para que comprendan lo que los posibles impactos sociales de un proyecto propuesto pueden significar para ellos.
- Mejorar el contenido local y los contratos de aprovisionamiento locales.
- Realizar monitoreo de cuestiones sociales.
- Diseñar e implementar medidas de inversión social.
- Establecer mecanismos de indemnización apropiados.
- Diseñar e implementar mecanismos de reclamos.
- Negociar acuerdos de impactos y beneficios (AIB).
- Elaborar planes de gestión de impacto social (PGIS).
- Preparar documentación para la entidad regulatoria.
- Identificar problemas y/u obligaciones para abordar la gestión de impactos sociales que se deban incluir en contratos con subcontratistas del proyecto.
- Preparar documentación de cumplimiento estándar sobre desempeño para una entidad financiera (Banco Mundial, CFI, otro banco de desarrollo multilateral o un banco adherido a los Principios de Ecuador).
- Elaborar una evaluación con la diligencia debida o auditoría de desempeño social para una comunidad, una ONG o una entidad financiera.
- Asistir en la planificación del cierre.

Fuente: AIEI (2015) con adapatación propia.

Cómo establecer colaboraciones

Una evaluación es un equilibrio entre los conocimientos técnicos expertos y la independencia que aporta el equipo de investigación, y la relevancia de las políticas, la orientación estratégica y la coordinación operativa con las que contribuyen los responsables de las políticas y los encargados del programa en el equipo de políticas públicas. Se puede utilizar una gama de modelos para crear e implementar esta alianza entre los equipos de investigación y de políticas públicas. La elección de la modalidad dependerá del contexto y de los objetivos de la evaluación de impacto, teniendo en cuenta una gama de riesgos.

Por un lado, un equipo de investigación totalmente independiente, en colaboración limitada con el equipo de políticas públicas, puede generar una evaluación de impacto desvinculada de las preguntas de interés de políticas, o que use una metodología limitada por la falta de interacciones con los encargados del programa. Por otro lado, un equipo de investigación plenamente integrado con el equipo de políticas públicas puede crear riesgos de conflictos de interés, o conducir a la censura de ciertos resultados si no se aplican los principios de la ciencia abierta.

Además, las evaluaciones a menudo pueden tener múltiples objetivos, entre ellos construir capacidad de evaluación con los organismos del gobierno o sensibilizar a los operadores del programa ante las realidades de sus proyectos al llevarse a cabo en el terreno. Estos objetivos más amplios también pueden determinar parcialmente el modelo que se elegirá. En general, lo que más importa en la calidad de la evaluación de impacto es si el enfoque de asociación producirá estimaciones no sesgadas de los impactos del programa. Siempre que se respeten los principios éticos de la investigación y la ciencia abierta, la ausencia de sesgo y la objetividad tienden a ser más cruciales para la calidad de la evaluación de impacto que la independencia funcional de los equipos de investigación y de políticas. En la práctica, a menudo se requiere una estrecha colaboración entre ambos para asegurar la elaboración de una estrategia de evaluación de impacto

de la más alta calidad. Algunas modalidades de modelos de colaboración sugeribles, se aprecian en la **Tabla 10.6**

Tabla 10.6. Modelos de colaboración

El modelo de externalización
Para los encargados del programa, siempre atareados gestionando operaciones complejas, a menudo es atractivo contar con un equipo externo encargado de diseñar e implementar la evaluación de impacto. Los modelos de externalización pueden adoptar diferentes formas. Los administradores de programa a veces intentan externalizar el diseño de la evaluación de impacto, así como la realización de diversas encuestas (normalmente, una encuesta de línea de base y de seguimiento) con una sola entidad en el marco de un contrato amplio. En otros casos, los administradores de programa primero externalizan el diseño y siguen con contratos de diversas fases de la recopilación y del análisis de datos. Sin embargo, externalizar totalmente la evaluación de impacto puede implicar riesgos considerables. Se dan lo siguientes casos: • Se entrega al equipo contratado un conjunto de parámetros del programa previamente definidos, con escaso margen para debatir sobre los planes de diseño y de implementación, o sobre el alcance, para dar forma a la investigación. • En otros casos, puede que no están definidas las reglas del programa y las modalidades de implementación necesarias para diseñar una buena evaluación de impacto. En esos casos, el equipo contratado encargado de dicha evaluación tiene una influencia limitada para asegurar que se definan estos elementos. • En otros casos, puede que el programa ya haya sido diseñado o que la implementación haya comenzado, lo cual puede limitar seriamente las opciones metodológicas de la evaluación. Para mitigar estos riesgos, normalmente es preferible que el equipo de políticas de introducción de innovaciones ya tenga hecho un diseño de evaluación de impacto, que incluya una estrategia de identificación, indicadores de resultados clave cálculos de potencia iniciales y tamaños aproximados de la muestra. Esto podría realizarse a través de un comité de supervisión o mediante una revisión técnica y científica regular de los productos de la evaluación. En su conjunto, estas medidas de mitigación señalan que es probable que el modelo más efectivo no sea totalmente externalizado.

El modelo de alianza

Es posible establecer alianzas mutuamente beneficiosas cuando los investigadores tienen interés en llevar a cabo investigaciones sobre una pregunta de políticas de introducción de innovaciones, y cuando éstos y los asistentes del programa procuran asegurar que su proyecto cuente con una evaluación de impacto de buena calidad. Los investigadores tienen incentivos para abordar nuevas preguntas que se añadirán a la base de evidencia global, y para ampliar el alcance de la evaluación de impacto y contribuir a que sea más visible. El equipo de investigación puede movilizar parte del financiamiento para la evaluación de impacto si los objetivos de los financiadores están estrechamente alineados con el objeto de investigación de la evaluación. Presenta este modelo riesgos:

- Es posible que los investigadores procuren incorporar elementos novedosos en la investigación de la evaluación de impacto que quizá no están totalmente alineados con los objetivos inmediatos de las políticas a nivel local, aunque puedan añadir valor en t.rminos más globales.
- Por su parte, los responsables de las políticas de introducción de innovaciones y los asistentes del programa quizá no siempre sepan apreciar el rigor científico necesario para emprender evaluaciones de impacto rigurosas, y quizá tengan una mayor tolerancia que el equipo de investigación a los riesgos potenciales de la evaluación de impacto.

Para mitigar esos riesgos, los objetivos del equipo de investigación y de los equipos de pol.ticas de introducción de innovaciones deben estar estrechamente alineados.

El modelo plenamente integrado

Algunas evaluaciones de impacto se implementan en un modelo plenamente integrado donde los equipos de investigación y de implementación del programa son el mismo, y son responsables tanto de la investigación como de las funciones del programa. Los investigadores a veces adoptan este enfoque en los ensayos de eficacia, donde se prueban nuevas intervenciones para la prueba de concepto. En este caso, los investigadores generalmente prefieren mantener el control de la implementación para asegurar que el programa se ponga en marcha siguiendo el dise.o original lo más estrechamente posible. Si bien los resultados de estas evaluaciones de impacto tienen la mayor capacidad para probar las teorías fundamentales y para establecer si una determinada intervención puede funcionar en circunstancias ideales, el riesgo de este enfoque es que los resultados pueden tener una validez externa limitada.

Elección de un equipo de investigación como socio

Los responsables de las políticas de introducción de innovaciones y los encargados del programa también tienen que decidir con quien asociarse. Las preguntas clave son si el equipo de investigación, o partes del mismo, puede ser un equipo local, y qué tipo de ayuda externa se requerirá. La capacidad de investigación varía en gran medida de un país a otro. Tome en cuenta:

- A menudo se contrata a las empresas internacionales cuando se requieren habilidades concretas, y también pueden asociarse con empresas locales. Las funciones de recopilación de datos generalmente son gestionadas por estas últimas, debido a su profundo conocimiento del contexto y del entorno local.
- Hay una marcada tendencia mundial a asegurar la plena participación de los investigadores locales en la evaluación de impacto.
- A medida que aumenta la capacidad de evaluación, es más habitual que los gobiernos, las empresas privadas y las instituciones multilaterales implementen evaluaciones de impacto en asociación con equipos de investigación locales. La participación de los investigadores locales puede aportar un valor fundamental a la evaluación de impacto gracias a su conocimiento del contexto local. En algunos países, la autorización de la investigación se concede solo a los equipos que incluyen a investigadores locales. En general, el administrador de la evaluación es el que evalúa la capacidad local y determina quién será responsable de qué aspectos del trabajo de evaluación.

Otra pregunta es si trabajar con una empresa privada o con un organismo público, por lo que:

- Las empresas privadas o los institutos de investigación pueden ser más fiables para proporcionar resultados de manera oportuna pero, una vez que se ha firmado un contrato, *las empresas privadas a menudo están menos dispuestas a incorporar en la evaluación elementos que podrán encarecerla.*
- El equipo de investigación también puede trabajar con instituciones de investigación y universidades, cuya reputación y conocimientos técnicos expertos garanticen que las partes interesadas acepten los resultados de la evaluación. Sin embargo, en ocasiones esas instituciones *carecen de la experiencia operativa* o de la capacidad para ejecutar ciertos aspectos de la evaluación, como la recopilación de datos.
- Por lo tanto, es posible que sea necesario subcontratar algunos aspectos con otro socio.

El desarrollo de capacidades en el sector público también puede ser un objetivo y se puede incluir como parte de los términos de referencia de la evaluación de impacto. Cualquiera sea la combinación de contrapartes a la que finalmente se llegue, será esencial efectuar un análisis sólido de las actividades de evaluación de los colaboradores potenciales en el pasado para tomar una decisión bien fundamentada. Particularmente, cuando se trabaja con un organismo público con múltiples responsabilidades, la capacidad y disponibilidad de un equipo de investigación interno para emprender las actividades de evaluación de impacto tienen que ser estimadas a la luz de otras actividades por las que deben rendir cuentas. Es importante tener conciencia de la carga de trabajo para valorar no solo cómo influir en la calidad de la evaluación que se lleve a cabo, sino también en el costo de oportunidad de la evaluación con respecto a otras iniciativas de las cuales es responsable el organismo público.

Fuente: Gertler et al. (2017) con adaptación propia.

Determinando el cronograma

Una elección más amplia para generar grupos de comparación, se hace a través de la planificación previa que facilite la recopilación de datos de *línea de base*, ayudando a las partes interesadas a alcanzar un consenso a propósito de los objetivos del programa y de las preguntas de interés. En la etapa de diseño del proyecto, la planificación de las evaluaciones de forma temprana, debe programarse su ejecución para evaluar el programa una vez que alcance la madurez para ser estable.

Las reformas incipientes y/o los proyectos piloto suelen sufrir revisiones constantes tanto en su contenido como al cuándo, dónde y por quién serán implementados. Los diversos actores, como los proveedores del programa, necesitarán tiempo para adaptarse a las nuevas reglas operativas que demandan las evaluaciones y deberán ser claras para generar contrafactuales adecuados para ejecutarlas, cuando los programas estén bien establecidos. Otro aspecto clave de la programación, es cuánto tiempo se requiere antes de que los resultados se puedan medir.

Se debe ubicar un equilibrio adecuado que depende en gran parte del contexto; esto es, si se evalúa demasiado temprano, existe el riesgo de encontrar un impacto parcial o nulo; si se evalúa demasiado tarde, existe el riesgo de que el programa pierda el apoyo de los donantes y del público o de que se amplíe un programa mal diseñado. Así, a fin de lograr un cronograma de evaluación impacto, se sugiere seguir las recomendaciones de la **Tabla 10.7.**

Tabla 10.7. Recomendaciones para realizar cronograma de evaluación de impacto

Ciclo del programa, duración del programa, implementación, retrasos potenciales
La evaluación de impacto debe: • Ajustarse al ciclo de implementación del programa; • No puede impulsar el programa que se evalúa. • Estar sujetas a los plazos del programa • Deben alinearse a su duración prevista. • Deben adaptarse a los posibles desfases en la implementación cuando los programas tardan en asignar beneficios o se retrasan debido a factores externos. • En general, a pesar de que la programación en el tiempo de la evaluación debería incluirse en el proyecto desde el comienzo, los evaluadores deberían estar dispuestos a ser flexibles e introducir modificaciones a medida que se ejecuta el proyecto. • Además, deben adoptarse provisiones para dar seguimiento a las intervenciones, utilizando un sistema de monitoreo de modo que el trabajo de evaluación se fundamente en el progreso real de la intervención.
Tiempo previsto para que el programa influya en los resultados y su naturaleza La programación de la recopilación de los datos de seguimiento debe tener en cuenta: • Cuánto tiempo se requiere después de que se ejecute el programa para que los resultados se manifiesten. • La cadena de resultados del programa ayuda a identificar los indicadores de resultados y el momento adecuado para medirlos. Algunos programas (como los de apoyo al ingreso) procuran proporcionar beneficios a corto plazo, mientras que otros (como los de educación básica) procuran tener objetivos a más largo plazo. • Ciertos resultados tardan más en manifestarse (como los cambios en un proceso de introducción de innovaciones tecnológica o la educación por la introducción de un nuevo proceso de innovación basada en TI) que otros (como los ingresos provenientes de un programa de capacitación). Los datos de seguimiento se pueden: • Recopilar más de una vez, de modo que se puedan tener en cuenta y se puedan contrastar los resultados de corto y mediano plazo. • Al ser recogidos durante la implementación del programa quizá no capturen el pleno impacto del mismo si los indicadores se miden demasiado temprano. Aun así, es muy útil documentar los impactos de corto plazo, que también pueden proporcionar información acerca de los resultados previstos a más largo plazo, útiles para producir resultados tempranos de la evaluación de impacto, que pueden estimular el diálogo entre los equipos de investigación y de políticas de introducción de innovaciones, y mantener el contacto con la muestra de evaluación de modo de reducir el desgaste de la muestra a lo largo del tiempo. • En ocasiones, los impactos de largo plazo son objetivos explícitos del programa, pero es posible que incluso un diseño sólido de evaluación de impacto no pueda resistir al paso del tiempo. Las encuestas de seguimiento que miden los resultados de largo plazo después de implementar el programa a menudo producen la evidencia más convincente en lo que se refiere a la efectividad del programa.

Ciclos de elaboración de políticas de introducción de innovaciones.
La programación de una evaluación también debe tener en cuenta cuando se requiere cierta información para fundamentar las decisiones de políticas y debe sincronizar las actividades de evaluación y de recopilación de datos con momentos clave en la toma de decisiones. La producción de resultados debería programarse para fundamentar los presupuestos, la ampliación del programa u otras decisiones de políticas.

Fuente: Gertler et al. (2017) con adaptación propia.

Determinando el presupuesto

El presupuesto constituye uno de los últimos pasos para hacer operativo el diseño de la evaluación. En este apartado, se ofrecen algunos datos sobre el costo de una evaluación de impacto, se debate cómo elaborar un presupuesto para una evaluación y se sugieren algunas opciones de financiamiento. **Ver Tabla 10.8.**

Tabla 10.8. Recomendaciones para determinar el presupuesto de evaluación de impacto

Análisis de los datos de costos No es sorprendente observar costos directos de las actividades de la evaluación que oscilan entre US$130.000 y US$2.78 millones, con un costos promedios cercano a US$1 millón. Aunque estos costos varían en gran medida y pueden parecer elevados en términos absolutos, se debe tomar en cuenta que: • Las evaluaciones de impacto suelen constituir solo un pequeño porcentaje de los presupuestos generales del programa. • Además, el costo de llevar a cabo una evaluación de impacto debe compararse con los costos de oportunidad de no efectuar una evaluación rigurosa y, por lo tanto, de implementar potencialmente un programa inefectivo. • Las evaluaciones permiten a los investigadores y a los responsables de las políticas de introducción de innovaciones, identificar qué programas o características del programa funcionan, cuáles no funcionan y qué estrategias pueden ser las más efectivas y eficientes para alcanzar los objetivos del programa. • En este sentido, los recursos necesarios para implementar una evaluación de impacto constituyen una inversión relativamente pequeña pero importante. • En estas evaluaciones, como en casi todas en las que no se pueden usar los datos existentes, *el costo más importante corresponde a la recopilación de nuevos datos*, que puede llegar en promedio, *al 65% del costo de la evaluación*. • Estas cifras reflejan diferentes tamaños y tipos de evaluaciones. El costo relativo de la evaluación de un programa piloto suele ser superior al costo relativo de la evaluación de un programa a nivel nacional o universal.

- Además, algunas evaluaciones solo requieren una encuesta de seguimiento o pueden usar las fuentes de datos existentes, mientras que otras necesitan llevar a cabo múltiples rondas de recopilación de datos.
- Los costos de recopilación de datos dependen sobre todo de las capacidades del equipo local, de los recursos disponibles y de la duración del trabajo de campo. Para saber más sobre cómo determinar los costos de una encuesta en un contexto particular, se recomienda al equipo de evaluación que primero entre en contacto con el organismo nacional de estadística y que busque información entre los equipos que hayan llevado a cabo un trabajo de encuestas en el país.

Elaboración de un presupuesto

Se necesitan muchos recursos para implementar una evaluación de impacto rigurosa, sobre todo cuando se trata de recopilar datos primarios. Las partidas presupuestarias incluyen:

- Los honorarios para al menos un investigador principal, un asistente de investigación, un coordinador del trabajo de campo, un experto en muestreo y los encuestadores.
- También se debe considerar el tiempo del personal del proyecto para proporcionar orientación y apoyo a lo largo de la evaluación. Estos recursos humanos pueden consistir en investigadores y expertos técnicos de organizaciones internacionales, consultores internacionales o locales y personal del programa local. Los costos de viaje y viáticos también se deben presupuestar.
- También se deben considerar en la planificación los recursos para la divulgación, con frecuencia en forma de talleres, informes y documentos académicos.
- Como se ha señalado, el costo más importante suele ser el relacionado con la recopilación de datos (que incluye la creación y la prueba piloto de la encuesta), los materiales y el equipo para recoger los datos, la capacitación de los encuestadores, sus salarios, los veh.culos y el combustible, y las operaciones de digitación de datos.

Calcular todos estos costos requiere considerar algunos supuestos, por ejemplo, acerca del tiempo que llevaría completar el cuestionario y de la duración de los viajes entre los emplazamientos. Los costos de una evaluación de impacto pueden repartirse a lo largo de varios ejercicios fiscales. Una vez más, es probable que las demandas presupuestarias sean mayores durante los años en que se recopilan los datos.

Opciones para financiar las evaluaciones

El financiamiento de una evaluación puede provenir de numerosas fuentes, entre ellas:

- Recursos para el proyecto.
- Presupuestos directos del programa.
- Ayudas a la investigación o financiamiento de los donantes.

A menudo, los equipos de evaluación procuran tener una combinación de fuentes para generar los fondos necesarios. A pesar de que el financiamiento para las evaluaciones solía provenir sobre todo de presupuestos para la investigación, un énfasis creciente en la elaboración de políticas basadas en la evidencia ha aumentado el financiamiento proveniente de otras fuentes. En los casos en que es probable que una evaluación zanje una brecha de conocimientos considerable de interés para la comunidad de desarrollo en términos más amplios, y donde se pueda aplicar una evaluación creíble y robusta, se debería instar a los responsables de las políticas de introducción de innovaciones, a buscar financiamiento externo, debido al bien público que los resultados de la evaluación proporcionarán. Las fuentes de financiamiento son:

- El gobierno.
- Los bancos de desarrollo.
- Las organizaciones multilaterales.
- Los organismos de las Naciones Unidas.
- Las fundaciones
- Las instituciones filantrópicas
- Las organizaciones de investigación y evaluación, como la Iniciativa Internacional para la evaluación de Impacto.

Fuente: Gertler et al. (2017) con adaptación propia.

CAPÍTULO 11.

Evaluación de Impacto Social en la Era POSCOVID-19

La economía mundial vive una crisis sanitaria, humana y económica sin precedentes en el último siglo y que evoluciona continuamente. Ante la pandemia de la enfermedad por coronavirus (COVID-19), las economías se cierran y paralizan, y las sociedades entran en cuarentenas más o menos severas, medidas solo comparables a las de situaciones de guerra. Aunque no se sabe cuánto durará la crisis ni la forma que podría tener la recuperación, cuanto más rápida y contundente sea la respuesta, menores serán los efectos negativos. Algunos de los mecanismos tradicionales de mercado podrían no ser suficientes para enfrentarla debido a la interrupción de las actividades productivas y la consiguiente contracción de demanda.

El Estado debe asumir actividades de planificación que generen las condiciones para sostener y después estimular la oferta y la demanda. Esto requiere capacidades y recursos públicos, e instrumentos de política que deben diseñarse en línea con las capacidades productivas de cada país, buscando preservar las capacidades instaladas. Las medidas de política para enfrentar la crisis están en permanente cambio, como se ve en el Observatorio COVID-19 en América Latina y el Caribe, de la CEPAL, que actualiza esa información frecuentemente.

Más allá de cualquier modelo predictivo, la acción de los gobiernos se está llevando a cabo sobre la base de procesos de prueba y error. En algunos casos la crisis sanitaria está adquiriendo elementos de crisis política, como se ve en la contradicción frecuente entre autoridades locales y gobiernos centrales o entre países en el seno de bloques de integración regional. Más aun, la desprotección de los sectores más pobres y las dificultades que enfrentan para obtener bienes básicos imprescindibles ya ha llevado a estallidos sociales.

Impacto económico para América Latina y el Caribe

América Latina y el Caribe enfrenta la pandemia desde una posición más débil que la del resto del mundo. Antes de la pandemia, la CEPAL preveía que la región crecería un máximo del 1.3% en 2020. Sin embargo, los efectos de la crisis han llevado a cambiar esa previsión y pronosticar una caída del PIB de al menos un 1.8%. Sin embargo, no es de descartar que el desarrollo que la pandemia lleve a previsiones de contracciones de entre un 3% y un 4%, o incluso más. El impacto económico final dependerá de las medidas que se tomen a nivel nacional, regional y mundial.

El COVID-19 afecta a la región a través de cinco canales externos de transmisión **(ONU CEPAL, 2020a):**

- **La disminución de la actividad económica de sus principales socios comerciales y sus efectos.** La región depende marcadamente de sus exportaciones, cuyo volumen y valor se reducirán por la recesión mundial. La magnitud del impacto final dependerá de la estructura sectorial de cada país.
- **La caída de los precios de los productos primarios**. Las marcadas caídas de esos precios y el deterioro de los términos de intercambio tendrán fuertes efectos negativos en los niveles de ingreso de las economías latinoamericanas dependientes de esas exportaciones, aunque con diferencias significativas entre ellas. La contracción de la demanda mundial, en particular la de China, uno de los mayores consumidores e importadores de productos primarios, jugará un papel destacado en la disminución de sus precios. A esto se agrega que una crisis geopolítica en el mercado del petróleo llevó a una reducción del 24% de los precios en menos de una semana a principios de marzo de 2020.
- **La interrupción de las cadenas globales de valor.** La disrupción de las cadenas de suministro, comenzando por los proveedores chinos y luego por la producción europea y estadounidense, afectaría prin-

cipalmente a México y el Brasil, cuyos sectores manufactureros son los más grandes de la región.

- **La menor demanda de servicios de turismo.** En particular, los pequeños estados insulares en desarrollo (PEID) del Caribe pueden ser muy afectados. Si los mayores efectos del COVID-19 se sienten en el segundo trimestre de 2020 y se concretan en prohibiciones o autorrestricciones de viajar de tres meses de duración, la actividad turística en el Caribe se contraería un 25%.
- **La intensificación de la aversión al riesgo y el empeoramiento de las condiciones financieras mundiales**. Esto conlleva una mayor demanda de activos seguros (por ejemplo, las tasas de rendimiento de valores de los Estados Unidos han llegado a niveles históricamente bajos), una menor demanda de activos financieros de la región y una importante depreciación de las monedas de sus países, como está ocurriendo.

La pandemia del COVID-19 impactó América Latina y el Caribe en un momento de debilidad de su economía y de vulnerabilidad macroeconómica. En el decenio posterior a la crisis financiera mundial (2010-2019), la tasa de crecimiento del PIB regional disminuyó del 6% al 0.2%; más aún, el período 2014-2019 fue el de menor crecimiento desde la década de 1950 (0.4%). Será la causa de la mayor crisis económica y social de la región en décadas, con efectos muy negativos en el empleo, el combate a la pobreza y la reducción de la desigualdad **(ONU CEPAL, 2020b).**

Los sectores más afectados por las medidas de distanciamiento social y cuarentena son los de servicios, que, en gran medida, dependen de contactos interpersonales. En la región, los sectores que podrían sufrir las mayores contracciones(comercio, transporte, servicios empresariales y servicios sociale) proveen el 64% del empleo formal. Además, el 53% del empleo de la región se da en actividades informales, que serán significativamente afectadas por basarse principalmente en contactos interpersonales **(ONU CEPAL, 2020a).**

Impactos en el comercio internacional

De acuerdo a la **ONU CEPAL (2020a)**, los efectos del COVID-19 agravan las ya débiles perspectivas del comercio exterior de América Latina y el Caribe. El acuerdo de *fase 1* entre los Estados Unidos y China celebrado en enero, en el que China se comprometió a aumentar sus importaciones de bienes y servicios de los Estados Unidos por lo menos en 77.000 millones de dólares en 2020, puede desplazar las exportaciones latinoamericanas y caribeñas a China en las mismas categorías de productos. Junto con la marcada caída de los precios del petróleo ya mencionada, el COVID-19 llevará a una contracción del comercio internacional para la región en términos de valor y de volumen, a través de su impacto directo en la región y a través de sus socios comerciales.

Se suponen diferentes tasas de crecimiento del PIB en sus principales socios comerciales y diferentes reducciones de los precios en las exportaciones de esta. Se estima que el valor de las exportaciones de la región caerá por lo menos el 10.7% en 2020. La mayor parte de la reducción del valor de las exportaciones corresponde a la caída de sus precios, estimada en un 8.2%; además, se espera que el volumen exportado registre una contracción del 2.5%.

En el plano subregional, el mayor impacto lo sentirán los países de América del Sur, que se especializan en la exportación de bienes primarios y, por lo tanto, son más vulnerables a la disminución de sus precios. En cambio, el valor de las exportaciones de Centroamérica, el Caribe y México registraría una caída menor que el promedio de la región, debido a sus vínculos con los Estados Unidos y su menor exposición a la disminución de los precios de los productos primarios. Los países exportadores de petróleo experimentarán la mayor pérdida en el valor de ventas al exterior. México, la República Bolivariana de Venezuela, el Ecuador y Colombia pueden ser los países más afectados, ya que sus costos de producción son más altos que los de muchos otros productores y, por lo tanto, tienen menos capacidad para soportar un periodo prolongado de precios bajos.

La crisis del COVID-19 causó un impacto en el desempeño exportador de la región también por su efecto sobre las importaciones utilizadas para producir exportaciones. México y Chile son los países más expuestos a una caída de la oferta de China, que suministra alrededor del 7% de sus insumos intermedios. Le siguen Colombia y el Perú, que importan de China el 4.5% y el 5% de sus insumos intermedios, respectivamente.

México es el país más expuesto a los cambios en las condiciones de la oferta y la demanda en los Estados Unidos, especialmente en el sector manufacturero. Costa Rica también está muy expuesta a las condiciones económicas de los Estados Unidos, ya que alrededor del 10% de su PIB depende de la oferta y la demanda de ese país. Los países más expuestos a los cambios en las condiciones de la oferta y la demanda en la Unión Europea son Chile, México y el Brasil, ya que alrededor del 5% de su PIB depende del valor añadido de los sectores de servicios y manufacturas en ese mercado.

Acceso y uso de Internet

Las medidas para detener la propagación de coronavirus han acelerado el ritmo al que el trabajo y la educación pasan al ámbito digital **(ONU, CEPAL, 2020a).** Las tecnologías digitales han disminuido el impacto de la pandemia en algunas profesiones y en la educación, al tiempo que han permitido sostener comunicaciones personales y actividades de entretenimiento en los hogares. Aunque más del 67% de los habitantes de la región usaron internet en 2019 y la penetración de la banda ancha ha aumentado marcadamente, el aumento del uso de las tecnologías digitales puede exacerbar las desigualdades derivadas del distinto acceso a las mismas entre los países y entre los grupos de ingresos. Mientras en 2017 más del 80% de la población estaba conectada a Internet móvil en Chile, el Brasil, Costa Rica y el Uruguay, esa cifra se reducía al 30% en Guatemala, Honduras, Haití y Nicaragua.

También existen fuertes desigualdades en la tasa de conectividad según segmentos de ingresos. En América Latina y el Caribe, la brecha entre los quintiles más ricos y los más pobres es mayor en países como Honduras (58 puntos porcentuales) y el Perú (60 puntos porcentuales), y menor en países como Chile (22 puntos porcentuales) y el Uruguay (17 puntos porcentuales).

ALAyC: Oportunidades de evaluación de impacto social

Incluso antes de la difusión del COVID-19, la situación social en América Latina y el Caribe (ALAyC) se estaba deteriorando, como muestran el aumento de los índices de pobreza y de extrema pobreza, la persistencia de las desigualdades y el descontento generalizado. En ese contexto, la crisis tendrá repercusiones negativas en la salud y la educación, así como en el empleo y la pobreza, dinde la comunidad requiere de estudios de evaluación de impacto social en la introducción de innovaciones **(ONU CEPAL, 2020a):**

Los sistemas de salud

Habrá fuertes impactos en el sector de la salud por la escasez de mano de obra calificada y de suministros médicos, así como por los aumentos de los costos. La mayoría de los países no han invertido lo necesario en salud. El gasto público del gobierno central en el sector, que en 2018 se situaba en un 2.2% del PIB regional **(ONU CEPAL, 2019)** está lejos del 6% del PIB recomendado por la OPS para reducir las inequidades y aumentar la protección financiera en el marco del acceso y la cobertura universal. Los recursos adicionales podrían contribuir a fortalecer el primer nivel de atención, con énfasis en medidas de prevención **(OPS, 2019).**

La mayoría de los países de la región se caracteriza por tener sistemas de salud débiles y fragmentados, que no garantizan el acceso universal necesario para hacer frente a la crisis sanitaria del COVID-19:

- Generalmente los sistemas de salud se organizan en torno a servicios en el sector público para las personas de bajos ingresos, servicios del seguro social para los trabajadores formales y servicios privados para quienes puedan costearlos. De esta manera, los sistemas permanecen segregados y claramente desiguales al ofrecer servicios de distinta calidad a diferentes grupos poblacionales. Si bien se han emprendido reformas para reducir esta fragmentación y expandir acceso al sistema de salud, aun son insuficientes.
- Además, los sistemas de salud tienden a ser geográficamente centralizados, con servicios y médicos especializados concentrados en pocos centros urbanos.

Las instalaciones son insuficientes para el nivel de demanda previsto y dependen en gran medida de las importaciones de equipamiento e insumos. Este es un problema importante porque, al 11 de marzo de 2020, 24 países del mundo habían restringido las exportaciones de equipo médico, medicamentos o sus ingredientes **(*The Economist,* 2020).** En 2018, solo siete países de la región contaban con un número significativamente más alto de camas de hospital por cada 1.000 personas que el promedio mundial.

Hay grandes brechas en el acceso a los sistemas de salud. La participación en los planes de seguro de salud para las personas empleadas de 15 años o más era solo del 57.3% en 2016, y entre la población del decil de ingresos más bajos, la cobertura era solo del 34.2%. A ello se suma que las dificultades para acceder a los centros de salud son agudas en las zonas rurales y remotas.

Los sistemas de salud de varios países de la región ya estaban bajo presión a causa de la epidemia de dengue: en 2019 se infectaron más de 3 millones de personas (la mayor cifra registrada en la historia de la región) y 1.538 personas murieron a causa de la enfermedad. El Brasil tuvo el mayor número de casos: 2,2 millones de personas **(OPS, 2020).**

La población cubierta por seguros médicos privados podría tener que hacer frente a elevados copagos para acceder a las pruebas de coronavirus, lo que sería un obstáculo a la detección temprana. En 2016 el gasto en salud

de bolsillo de los hogares como proporción del gasto corriente total en salud en América Latina y el Caribe (37.6%) duplicó con creces el nivel de la Unión Europea (15.7%) **(OMS, 2017).**

Como la estructura demográfica de la región es bastante joven, es posible que el impacto general sea menor que en los países desarrollados. En promedio, solo el 10% de la población de América Latina y el Caribe (casi 58 millones de personas) tiene 65 años o más. Los países con una distribución de la población más sesgada hacia adultos mayores, como Barbados, Cuba, el Uruguay, Aruba y Chile, podrían sufrir una presión mayor en los sistemas de salud.

Educación

Al 20 de marzo de 2020, la Argentina, Bolivia (Estado Plurinacional de), Chile, Colombia, el Ecuador, El Salvador, Guatemala, Honduras, Jamaica, Panamá, el Paraguay, el Perú, Santa Lucía, Trinidad y Tobago, el Uruguay y Venezuela (República Bolivariana de) habían suspendido las clases en todos los niveles educativos. En el Brasil se habían aplicado cierres localizados de centros educativos.

La interrupción de las actividades en centros educativos tendrá efectos significativos en el aprendizaje, especialmente de los más vulnerables. Los centros educativos también proporcionan seguridad alimentaria y cuidado a muchos niños, lo que permite a los padres tener tiempo para trabajar. La suspensión de las clases tendrá un impacto más allá de la educación, en la nutrición, el cuidado y la participación de los padres (especialmente de las mujeres) en el mercado laboral.

Alrededor de 85 millones de niños y niñas de la región reciben un desayuno, un refrigerio o un almuerzo en la escuela **(FAO/PMA, 2019).** Por lo tanto, es importante asegurar la continuidad de los programas de alimentación escolar. Aunque se han hecho planes para promover el uso de dispositivos digitales en los sistemas educativos, muchas instituciones educativas no cuentan con la infraestructura de tecnologías digitales necesaria. Además, existen brechas en el acceso a las computadoras y a Internet en

los hogares. Los procesos de enseñanza y aprendizaje a distancia no están garantizados.

Además, existen disparidades de acceso a los dispositivos digitales y a Internet de banda ancha entre las poblaciones urbanas y rurales, entre los sexos, entre las poblaciones que hablan o no el idioma oficial (español o portugués), y entre las poblaciones con y sin discapacidades. América Latina se enfrenta a desafíos en la formación de los docentes en materia de TIC. Por ejemplo, en el Brasil en 2018, solo el 20% de los docentes participaron en un curso de educación continua para el uso de computadoras e Internet para la enseñanza. En cuanto al uso de Internet, el 16% informó que la utilizaba una o más veces al día; mientras que el 20% lo hacía una vez a la semana, y el 18% al menos una vez al mes **(CGIB, 2019).**

Empleo y pobreza

Dadas las desigualdades económicas y sociales de la región, los efectos del desempleo afectarán de manera desproporcionada a los pobres y a los estratos vulnerables de ingresos medios. Es probable que la crisis aumente el empleo informal como estrategia de supervivencia. En 2019 el 53.1% de los trabajadores de América Latina y el Caribe trabajaba en el sector informal **(OIT, 2020).** Es probable que las familias más pobres envíen a sus hijos al mercado de trabajo , lo que aumentará las tasas de trabajo infantil. La OIT estima que actualmente el 7.3% de los niños de 5 a 17 años (unos 10,5 millones de niños) de la región trabajan.

ONU CEPAL (2019) mostró que la pobreza en la región aumentó entre 2014 y 2018. Debido a los efectos directos e indirectos de la pandemia, es muy probable que las actuales tasas de pobreza extrema (11.0%) y pobreza (30.3%) aumenten aún más en el corto plazo. Si los efectos del COVID-19 llevan a la pérdida de ingresos del 5% de la población económicamente activa, la pobreza podría aumentar 3.5 puntos porcentuales, mientras que se prevé que la pobreza extrema aumente 2.3 puntos porcentuales. Mayores deterioros de los ingresos implicarán aumentos aún mayores de la pobreza.

Impacto económico en las micro, pequeñas y medianas empresas

Dadas las desigualdades económicas y sociales de la región, los efectos del desempleo afectarán de manera desproporcionada a los pobres y a los estratos vulnerables de ingresos medios. Es probable que la crisis aumente el empleo informal como estrategia de supervivencia. En 2016 el 53.1% de los trabajadores de América Latina y el Caribe trabajaba en el sector informal **(OIT, 2020).** Es probable que las familias más pobres envíen a sus hijos al mercado de trabajo, lo que aumentará las tasas de trabajo infantil. La OIT estima que actualmente el 7.3% de los niños de 5 a 17 años (unos 10.5 millones de niños) de la región trabajan. **ONU CEPAL (2019)** mostró que la pobreza en la región aumentó entre 2014 y 2018. Debido a los efectos directos e indirectos de la pandemia, es muy probable que las actuales tasas de pobreza extrema (11.0%) y pobreza (30.3%) aumenten aún más en el corto plazo.

El futuro mundial posCOVID-19

De acuedo al la **ONU CEPAL (2020b),** se tiene:

- La crisis que sufre la región en 2020, con una caída del PIB del 5.3%, será la peor en toda su historia. Para encontrar una contracción de magnitud comparable hace falta retroceder hasta la Gran Depresión de 1930 (-5%) o más aún hasta 1914 (-4.9%).
- Los efectos de mediano plazo en materia de reorganización de la producción y del comercio internacional en términos de localización y tecnológicos son importantes. Los escenarios posibles que aún están abiertos son por lo menos tres: continuación de la globalización, pero sobre la base de nuevos modelos de gobernanza más receptivos al multilateralismo y la corrección de las desigualdades entre países; soluciones de alcance exclusivamente nacional, o una acentuación de la regionalización.

- Para la gran mayoría de los países de América Latina y el Caribe, las soluciones de alcance exclusivamente nacional no serían viables por razones de economías de escala, tecnológicas y de aprendizaje.
- Es posible que la mejor solución sea una nueva globalización con una gobernanza proclive a la inclusión y la sostenibilidad, pero para participar activamente en esa nueva globalización, América Latina y el Caribe debe integrarse productiva, comercial y socialmente. Para ello, la coordinación de nuestros países en materia macroeconómica y productiva es crucial para negociar las condiciones de la nueva normalidad, particularmente en una dimensión urgente en la actual crisis y en el mediano plazo: la del financiamiento para un nuevo estilo de desarrollo con igualdad y sostenibilidad ambiental.

El impacto social del COVID-19 en México

La contingencia sanitaria internacional provocada por la enfermedad COVID-19 amenaza la vida de millones de personas en el mundo. Aunque es una crisis sanitaria, la pandemia también tendrá un inevitable impacto en la economía, el comercio, los empleos, el bienestar y las condiciones de vida de la población. Para México, esto representa un reto de enorme complejidad por las múltiples implicaciones de la crisis y las fragilidades estructurales con las que se enfrenta.

La información que genera el Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) en materia de pobreza e ingreso, la evidencia disponible generada por ejercicios rigurosos de evaluación y los aprendizajes acumulados en materia de desarrollo social, permiten a la institución aportar elementos para dimensionar el problema que enfrenta el país, identificar alternativas e instrumentos para priorizar la protección social de los más vulnerables y sopesar algunas implicaciones fiscales, normativas y operativas de los cursos de acción que aquí se presentan.

Una primera aproximación a los efectos potenciales de la coyuntura anticipa aumentos preocupantes en la pobreza por ingresos, la pobreza laboral y el riesgo de que, en ausencia de políticas públicas orientadas a esta población, grupos de ingreso medio enfrenten condiciones de pobreza.

Considerando la decisión del Gobierno Federal, del 23 de abril de 2020, de incluir a 19 programas y acciones de desarrollo social en el conjunto de 38 programas prioritarios para atender la crisis **(DOF, 2020),** el CONEVAL analizó el nivel de incidencia que podrían tener estos 19 programas para apoyar a los sectores de la población que se considera serán los más afectados por la pandemia y las medidas para mitigar sus efectos: la población en condiciones de pobreza (urbana y rural), los trabajadores tanto del sector formal como del informal, las personas desempleadas y las Micro, Pequeñas y Medianas Empresas (MiPyME).

Para este análisis se calcularon los recursos financieros necesarios, como porcentaje del Producto Interno Bruto (PIB), para utilizar los programas existentes para canalizar medidas emergentes en dos escenarios de ampliación de cobertura: la atención de los programas al resto de su población potencial (población elegible pero aún no atendida en 2019) y 30% de incremento en su cobertura. Igualmente, se realizó una valoración de la estructura operativa existente y los desafíos operativos que enfrentan los programas.

La conclusión general es que esta crisis amenaza los avances de México en materia de desarrollo social y que afectará en mayor proporción a los grupos más vulnerables. Ante este desafío se impone la necesidad de ampliar y fortalecer las medidas emergentes de respuesta desplegadas a partir de los programas prioritarios, mejorar su capacidad operativa, asegurar que su focalización sea adecuada, así como considerar medidas adicionales de atención a quienes pierdan su fuente de ingreso como resultado de la contingencia y a sectores urbanos que están más expuestos.

La pandemia de COVID-19 nos enfrenta a la urgente necesidad de promover la construcción de un sistema de protección social que asegure la resiliencia y el avance progresivo de los derechos sociales en México,

siendo una oportunidad para realizar la evaluación de impacto en la introducción de innovaciones, lo mostrado en la **Tabla 11.1**

Tabla 11.1. Oportunidad de la evaluación de impacto social para introducción de innovaciones en México posCOVID

Pobreza y vulnerabilidad en México
Las diferentes crisis econ.micas que México ha enfrentado en el pasado han puesto en evidencia que, a pesar de ser un paÍs de ingreso medio, los altos niveles de pobreza y desigualdad acentúan la vulnerabilidad social frente a choques econ.micos adversos, afectando casi invariablemente a la población históricamente en desventaja. Los hogares mexicanos dependen fuertemente de sus ingresos laborales (el ingreso por trabajo subordinado e independiente representa más de 50% del ingreso corriente total de los hogares en cada uno de los deciles de ingreso) y cuatro de cada diez personas en México se encuentran en situaci.n de pobreza (41.9 % de la población). Entre 2008 y 2018, México logró una disminución de la pobreza de 2.5 %, al pasar de 44.4% a 41.9%. Las carencias sociales que presentaron mayores reducciones fueron el acceso a los servicios de salud, pasando de 38.4 % en 2008 a 16.2 % en 2018, así como la calidad y espacios de la vivienda, que en 2008 era de 17.7 % y en 2018 fue de 11.1%. En 2020, México enfrenta esta crisis en condiciones de vulnerabilidad. Entre otros factores, destacan la alta prevalencia de diabetes y enfermedades cardiovasculares, la precariedad laboral, problemas de acceso al agua y de hacinamiento que impiden la adopción generalizada de medidas preventivas y múltiples brechas de acceso a derechos sociales: salud, alimentación, educación, vivienda y seguridad social. Aunque esta crisis supone riesgos sanitarios y económicos inéditos, las afectaciones derivadas de las crisis de 1994-1996 y 2008-2010 muestran como la población se ve afectada por el aumento en el desempleo y la disminución de los ingresos y que los grupos de población que enfrenta mayores adversidades suelen ser los de menores ingresos.

¿Qué efectos podría traer la crisis producida por COVID-19 en la pobreza?

Algunos pronósticos para el país estiman una caída del PIB de entre 2 y 6 por ciento. En el caso del empleo, aunque México se ha caracterizado por tener una tasa de desempleo relativamente baja, las cifras de los primeros diagnósticos coinciden en que habrá pérdidas importantes, pues a principios de abril ya se habían perdido 346,000 empleos formales **(PDR, 2020).** En este sentido, El CONEVAL estimó los efectos potenciales que la actual coyuntura podría generar en los niveles de pobreza por ingresos de la población. Para este análisis se hicieron estimaciones a partir de las fuentes estadísticas utilizadas para medir la pobreza, la Encuesta Nacional de Ingreso y Gasto en los Hogares (ENIGH) y la pobreza laboral en el corto plazo con la Encuesta Nacional de Ocupación y Empleo (ENOE). Para la primera se construyeron dos escenarios, en función de la distribución del impacto entre diversos grupos de población. En el primer escenario se simula una caída generalizada en el ingreso equivalente a 5% ; en el segundo, una reducción en el ingreso más pronunciada para los hogares en pobreza urbana. En cada escenario se recalculó la pobreza por ingresos de acuerdo con la metodología del CONEVAL, ajustando las líneas vigentes en marzo de 2020 con las expectativas de inflación del Banco de México a agosto del 2020. Los resultados permiten identificar lo siguiente:

- En ambos escenarios, el total de personas en situación de pobreza por ingresos se incrementa, entre 7.2 y 7.9 %(entre 8.9 y 9.8 millones de personas).
- El número total de personas en situación de pobreza extrema por ingresos se incrementa entre 4.9 y 8.5 % (6.1 y 10.7 millones de personas).
- Sin políticas públicas que atiendan a la población con ingreso medio la cantidad de personas en situación de pobreza por ingreso puede aumentar.
- La crisis puede cambiar las condiciones de ingresos de la población.

La estimación del efecto en la pobreza laboral (población que no podría adquirir una canasta alimentaria con su ingreso laboral) se basó en la tendencia observada en la crisis de 2008-2009. Se estima un aumento en la tasa de desempleo de 3.3 a 5.3 % y de 37.3 a 45.8 % en la pobreza laboral en el segundo trimestre del 2020. Este aumento del Índice de Tendencia Laboral de la Pobreza (El ITLP es un indicador que el CONEVAL presenta de manera trimestral con base en la Encuesta Nacional de Ocupación y Empleo (ENOE), el cual permite identificar, de manera oportuna, la tendencia de los ingresos laborales de la población respecto a su poder adquisitivo de la canasta alimentaria) representaría la cifra más alta de este indicador desde 2005. Además de los posibles efectos en la pobreza, es necesario considerar que un sector de la población que deberá afrontar esta pandemia con mayores desventajas son las mujeres. Se estima que las mujeres representan 72.% del total de personas ocupadas en los sistemas de salud en América Latina, para las cuales las condiciones precarias de trabajo que se vuelven extremas con el aumento de horas laborales y el riesgo de contagio de COVID-19 **(ONU CEPAL, 2020a),** además de los costos y esfuerzo de participar en mayor proporción en las tareas de cuidado en casa y de trabajo doméstico. Ante estos escenarios, es necesario diseñar una estrategia para afrontar los efectos que esta emergencia detonará para los grupos más vulnerables en las dimensiones de bienestar económico y ejercicio de derechos sociales.

Al mismo tiempo, esta situación obliga a una reflexión seria para impulsar medidas de mediano y largo plazo que son imprescindibles para construir un sistema de protección social con enfoque de derechos que sea resiliente a los eventos críticos. Entre estas medidas destacan la implementación de nuevas estrategias para proteger a la población más vulnerable, consolidando la protección social a través de una pensión universal mínima de retiro e invalidez financiada con impuestos generales. También se encuentra la adopción de medidas de largo plazo que garanticen la seguridad ante riesgos, como el seguro de desempleo contributivo, la renta básica ciudadana o un piso mínimo solidario. Si la situación se enfrenta de manera proactiva, como una oportunidad para la innovación en protección social, es posible enfrentar esta crisis con mejores capacidades para la implementación de la política de desarrollo social.
Bienestar económico La pérdida de empleos traerá como consecuencia no solo la disminución en los ingresos de los hogares, sino la pérdida de las prestaciones laborales que el empleo formal brinda. De acuerdo con la medición multidimensional de la pobreza 2018, la carencia por acceso a la seguridad social fue la más alta, lo que significa que más de la mitad (57.3%) de la población no cuenta con seguridad social. Las personas más vulnerables en este sentido son las que se encuentran ocupadas en: i) trabajo subordinado, ii) trabajo por cuenta propia, iii) trabajo en los sectores más afectados y iv) ocupados sin seguro de desempleo. En este sentido, se podrían tomar las siguientes medidas específicas para disminuir el impacto de la emergencia sanitaria para las poblaciones que a continuación se mencionan:
Empleo formal Además de los apoyos que ha dado a conocer el Gobierno Federal para microempresas del sector formal e informal de la economía, el Estado podríaa complementar estas medidas mediante apoyos a la pequeña y mediana empresa, a través de créditos para proteger el empleo y de subsidios al empleo parcial o total de las cuotas obreropatronales que los empleadores deben asumir durante los meses de confinamiento. Con estas medidas se apoyaría a 43.8 % la población ocupada.
Empleo informal Se puede apoyar a este sector de la población a través de créditos, como en el caso del Programa Microcréditos para el Bienestar que otorga financiamiento y capacidades técnicas para el inicio o consolidación de micronegocios. En este contexto, el programa podría extender los periodos de vencimiento de créditos, financiar temporalmente el consumo privado y analizar la pertinencia del apoyo de la estrategia *Mes Trece* del Programa Jóvenes Construyendo el Futuro. Asimismo, se podrían incluir medidas que incentiven el empleo temporal. Lo anterior permitiría apoyar a 52% de la población ocupada.

Ingreso de las personas y los hogares
Se requieren medidas coordinadas de los tres órdenes de gobierno para promover que los hogares que tendrán dificultades en el ingreso puedan atender las necesidades más básicas; por ejemplo, a través de subsidios temporales en el costo de los servicios básicos de la vivienda. Asimismo, se debe asegurar la adecuada operación de los programas federales que realizan transferencias monetarias directas.
Ingreso y trabajo
Se debe analizar la implementación de nuevas estrategias para proteger a la población más vulnerable, por ejemplo, consolidar la protección social a través de una pensión universal mínima de retiro e invalidez financiada con impuestos generales. Ante la complejidad de identificar y transferir recursos oportunamente a cada uno de los grupos vulnerables afectado por esta crisis, y las limitaciones de los instrumentos de transferencias disponibles con los que actualmente cuenta el gobierno federal y los gobiernos locales, los cuales fueron diseñados para ofrecer protección social al ingreso de estas poblaciones (como se expone más adelante), no se deben descartar otros instrumentos posibles, como una transferencia básica en el periodo de crisis, no necesariamente universal, pero con una cobertura amplia de los trabajadores afectados.
Alimentación
En el corto plazo se deben tomar en cuenta los recursos materiales, económicos y humanos existentes para proveer a la población insumos necesarios para contar con una alimentación nutritiva, suficiente y de calidad. Para ello, podrían adoptarse las siguientes medidas: • Garantizar la operación del Programa de Abasto Social de Leche (LICONSA) y del Programa Abasto Rural (DICONSA). • Fortalecer el Programa de Precios de Garantía. • Impulsar la producción agropecuaria en el país a través del Programa Especial Concurrente para el desarrollo Rural Sustentable (PEC). • Dar seguimiento al comportamiento del mercado para detectar a tiempo alzas injustificadas en los precios de productos a través de la Procuraduría Federal del Consumidor (PROFECO). • Otorgar despensas a las familias de los menores que son beneficiados con la provisión de alimentos en las escuelas, en tanto las clases se encuentran suspendidas, a través, por ejemplo, del Fondo de Aportaciones Múltiples en su componente Asistencia Social (FAM-AS). Este fondo cuenta con recursos que financian la Estrategia Integral de Asistencia Alimentaria, que contempla la entrega de despensas a la población más vulnerable (niñas, niños, mujeres embarazadas, personas con discapacidad y adultos mayores). • Otorgar cupones para la adquisición de alimentos durante la contingencia sanitaria, como mecanismo para mejorar la seguridad alimentaria de los hogares.

Salud
Es fundamental asegurar que el sistema de salud funcione adecuadamente y se apliquen correctamente las medidas de seguridad sanitaria emitidas por el Gobierno Federal. En este sentido, el Estado mexicano debe preparar al sistema de salud para enfrentar de manera efectiva el incremento en la demanda de los servicios de atención médica. Se consideran pertinentes las siguientes medidas: • Fortalecer los servicios estatales y municipales de salud, mediante la dispersiónde recursos. • Continuar con la implementación de servicios, protocolos y sistemas de información homologados para todas las instancias de salud durante y después de la emergencia sanitaria. • Aumentar la disponibilidad de camas para hospitalización previendo el aumento en la demanda de estos servicios. • Agilizar la contratación de profesionales de la salud para trabajar durante la emergencia sanitaria. • Difundir y reforzar las acciones implementadas para mitigar los efectos negativos en la salud mental de la población y la atención a mujeres víctimas de violencia de género y violencia sexual. • En el mediano plazo, mejorar la capacidad resolutiva de los diferentes niveles de atención, a través de: 1) reforzamiento del equipo básico en el primer nivel de atención, 2) reorganización y fortalecimiento de la capacidad resolutiva de instalaciones de todos los niveles de atención.
Educación
Desde la suspensión de actividades presenciales en el sistema educativo, la Secretaría de Educación Pública (SEP) ha implementado diversas acciones para continuar las actividades de aprendizaje a distancia. En ese marco, se presentan las siguientes consideraciones: • La educación básica a distancia contempla habilidades y herramientas tecnológicas que ponen en desventaja a quienes forman parte de los hogares más pobres, lo cual puede profundizar la brecha educativa. • El aprendizaje mediante educación básica a distancia presenta dificultades para dar seguimiento y conocer su efectividad, lo cual puede incrementar la brecha educativa entre instituciones públicas y privadas y entre contextos rurales y urbanos. • La educación básica a distancia requiere del apoyo de los padres a los hijos, el cual está supeditado a la ocupación, disponibilidad de tiempo y su propia escolaridad, entre otros factores. • No se encontró información sobre modalidades educativas para diferentes contextos, como la indígena, comunitaria o migrante. • Es necesario implementar mecanismos de regularización al término de la emergencia sanitaria para homologar los conocimientos adquiridos por los alumnos.

¿Cómo llegar a la población más vulnerable?

Dado que los recursos son escasos y con la finalidad de destinar una mayor proporción del presupuesto público a mitigar los efectos de la crisis en los grupos más vulnerables de la población, es fundamental que la identificación de sus beneficiarios sea la adecuada en este contexto, privilegiando los recursos destinados a los programas identificados relevantes para atender pobreza. Para ello, consolidar un padrón único de beneficiarios es impostergable, éste puede ser construido a través de los padrones o las bases de datos de los programas prioritarios.

Fuente: Coneval (2020).

Estructura operativa estrategias ante COVID-19

Las fuentes del financiamiento para la expansión de la oferta gubernamental requerida en el esfuerzo por mitigar los efectos adversos de esta contingencia sanitaria son un elemento fundamental que debe formar parte de la discusión de expertos y funcionarios gubernamentales.

El 23 de abril de 2020 el Gobierno Federal estableció como prioritarios 38 programas, de los cuales 19 corresponden al ámbito de desarrollo social **(DOF, 2020).** Considerando lo anterior, se analiza el nivel de incidencia que pueden tener estos 19 programas y acciones de desarrollo social para apoyar a los sectores de la población que se considera serán los más afectados por la pandemia: la población en situación de pobreza (urbana y rural), los trabajadores del sector formal e informal, las personas desempleadas y las Micro, Pequeñas y Medianas Empresas (MiPyME). Al respecto, destaca lo siguiente:

- Ninguno de los programas se dirige explícitamente a la población en situación de pobreza, de acuerdo con la medición del CONEVAL, aunque podría coincidir con la priorización territorial aplicada por varios de los programas.
- Se espera que los mayores efectos adversos de la pandemia sean en las zonas urbanas.

- Hasta el momento ninguna acción está dirigida a la población que perdió su empleo a raíz de la emergencia sanitaria.
- Algunas de las medidas anunciadas por el Gobierno Federal, podrían ser insuficientes para llegar a toda la población que se verá afectada.

Para hacer una aproximación de los recursos necesarios para financiar acciones emergentes de los programas existentes se presentan dos escenarios. El objetivo de este ejercicio es estimar cuántos recursos necesitarían los programas para atender un incremento en su cobertura, atendiendo, por un lado, al resto de su población objetivo que cuenta con los requisitos de elegibilidad pero que no fueron atendidos por el programa en 2019, y, por otro lado, atendiendo un incremento en su cobertura de 30%. El ejercicio destaca lo siguiente:

- **Escenario 1.** Contemplando una cobertura adicional de 8,873,924 beneficiarios, se necesitarían $137,288.1 millones de pesos adicionales al presupuesto ejercido, por lo que se requeriría un presupuesto de $293.816.4 millones de pesos lo que representa 1.21 % del PIB.
- **Escenario 2**. Con el incremento de la cobertura hasta en un 30 %, se apoyaría a 2,815,133 beneficiarios, requiriendo $41,681.1 millones de pesos adicionales al presupuesto ejercido, por lo que el presupuesto total requerido sería de $198,209.4 millones de pesos, lo que representa 0.82 % del PIB.

Es necesario considerar que, además de la carga fiscal y el costo de oportunidad de ampliar las coberturas de los programas, también supone una carga operativa importante, lo que significa también una amplia movilización en campo del personal operativo en la mayoría de los programas. Además de lo anterior, para ampliar la cobertura resulta necesario que los programas cuenten con información de personas que solicitaron ser beneficiarias y que cumplen con los requisitos pero que no fueron beneficiarios. Las transferencias monetarias que entreguen los programas son necesarias, pero dados los cambios que la pandemia tendrá en el contexto

social, es indispensable considerarlas como medidas temporales e impulsar la reflexión sobre la nueva complejidad de los problemas que enfrentará la población y por ende los nuevos programas públicos requeridos para atender estos de manera efectiva.

ANEXO.
Introducción a STATA

Esta sección está diseñada para conocer los recursos de **STATA** https://www.stata.com/. Ver **imagen 1.**

Imagen 1. Sitio web de STATA

Fuente: sitio web STATA.

Con el fin de manipular los archivos del caso de estímulos a la innovación, adjuntos a la obra, utilizando **STATA 14.1** (en adelante **STATA**). Es un caso hipotético de información recolectada de pymes de base tecnológica (**pbt**) correspondiente al sector de desarrollo de software de la industria electrónica, en varias regiones de un país, con observaciones de año 1 a año 8.

Estructura de archivos (*File Structure*)

Estos ejercicios usan y generan muchos archivos. Existen principalmente tres tipos de archivos **STATA**:

- Los que contienen el conjuntos de datos, identificados por el sufijo .**dta**
- Los que contienen los programas **STATA**, identificados por el sufijo **.do**, y
- Los que contienen el registro y salida del trabajo hecho en **STATA** identificados por el sufijo **.log**.

Para mantener estos archivos organizados, se sugiere la siguiente estructura de directorios, ubicados en un disco c, de un equipo Macintosh:

"/Users/DCA/Desktop/eval/data"
"/Users/DCA/Desktop/eval/do"
"/Users/DCA/Desktop/eval/log"

Descripción de archivos y ejemplos en STATA

Los archivos de datos, son localizados en "/Users/DCA/Desktop/eval/data", del disco c de un equipo Macintosh, con tres archivos:

- **pei_01.dta.** Este archivo comprende los datos de pymes de base tecnológica (**pbt**) del año **1** que con **826** registros con **24** variables de información.
  1.ID
  2. year.Year of observation
  3. region. Region
  4. cluster. Cluster

5. ageceo. Age of CEO
6. genceo.Gender of CEO
7. educeo. Education of CEO
8. smesize. Sme size based on quantity of personnel
9. smefac. Facilities, resources for innovations
10. incinnov. Incentives for innovations per year
11. expopn. Sme operation expenditure per year
12. expinnov. Sme innovation expenditure per year
13.exptot. Sme total expenditure per year
14. dmmfd. Innovation male microcredit participant: 1= Y; 0=N
15. dfmfd. Innovation female microcredit participant: 1= Y; 0=N
16. weight. Innovation sampling weight
17. bdbaccess. Broadband access
18. pcirr. Proportion of region with internet broadband
19. raw1. Raw 1
20. raw2. Raw 2
21. raw3. Raw 3
22. raw4. Raw 4
23. raw5.Raw 5
24. raw6.Raw 6

- **pei_08.dta.** Este archivo es la versión levantada al **año 8** de la versión **pei_01.dta**. Incluye **303** nuevos pymes de base tecnológica (**pbt**), haciendo el número total **1,129** pymes.
- **pei_0108.dta.** Este es un conjunto de datos de panel restringido a **826** pymes de base tecnológica (**pbt**) entrevistados en los dos años. en formato de serie temporal.

El subdirectorio **.do** tiene los *archivos del programa* (**.do**) específicos para las diferentes técnicas de evaluación de impact. Estos archivos contienen todo el código **STATA** necesario para implementar los ejemplos de los capítulos aquí contenidos.

El subdirectorio **.log** contiene todas las salidas generadas al ejecutar los archivos **.do.**

STATA es un paquete de software estadístico que ofrece una gran cantidad de funciones estadísticas y procedimientos de estimación econométrica. Iniciar una sesión de **STATA** se realiza oprimiendo doble en el icono instalado en su escritorio. Ver **imagen 2.**

Imagen 2. Icono de STATA
instalado en el escritorio de un equipo Macintosh



El **STATA** El entorno informático consta de cuatro ventanas principales. El tamaño y la forma de estos las ventanas pueden cambiarse y moverse en la pantalla. En su versión **STATA 14**. Ver **Imagen A.2.**

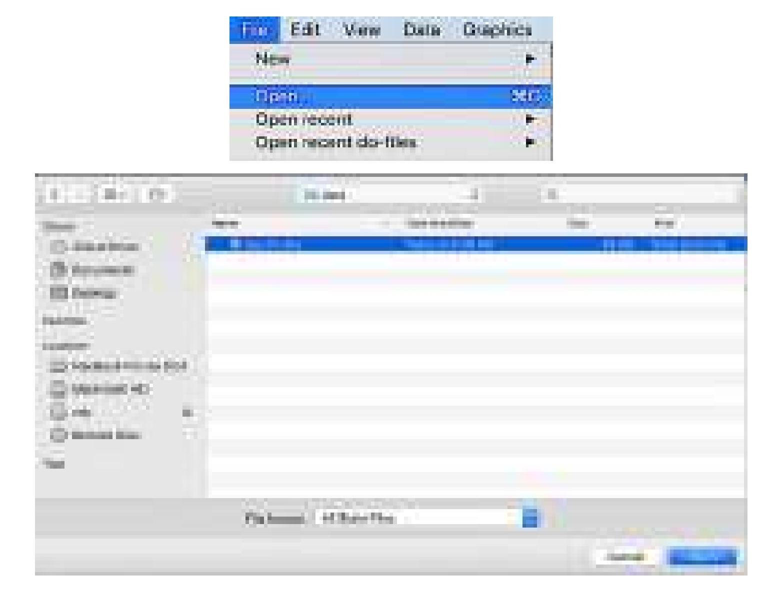
Imagen 2. Vista de STATA 14

Fuente: STATA 14.1.

Además de estas ventanas, el entorno **STATA** tiene un *menú* y una *barra de herramientas* en la parte superior (para realizar operaciones **STATA**) y una *barra de estado del directorio* en la parte inferior (que muestra el directorio actual). Es posible, usar el *menú y la barra de herramientas* para emitir diferentes comandos **STATA** (como abrir y guardar archivos de datos), aunque la mayoría de las veces se usa la *ventana de comandos* **STATA** para ejecutar programas de aplicación. Si está creando un registro archivo (discutido con más detalle más adelante), el contenido se puede mostrar en la pantalla, lo cual es útil si desea regresar y ver resultados anteriores de la sesión actual. Aunque no es el propósito de dar una capacitación en **STATA**, sí es la intención de dar a conocer al lector de sus principales características y forma de usarse a través de una serie de ejercicios prácticos, para su mayor claridad.

Ejemplo 1. Abriendo el conjunto de datos (*Open data set*)

Se hace al ingresar el siguiente comando en la *ventana de comandos* **STATA 14.1. File→Open→pei_01.dta**





La línea que se observa, es producto de la instrucción de apertura de archivo y con la ausencia de un mensaje de error en una segunda línea, implica que el comando se ha ejecutado con éxito. De ahora en adelante, solo se mostrará la *ventana de resultados* **STATA**. El seguimiento los puntos deben tenerse en cuenta:

- **STATA** asume que el archivo está en su formato con una extensión **.dta.** Por lo tanto, escribiendo **pei_08** es lo mismo que escribir **pei_08.dta.**
- Solo se puede abrir un conjunto de datos a la vez en **STATA**. En este caso, por ejemplo, si abre otro conjunto de datos **pei_01.dta**, reemplazará **pei_08.dta** con **pei_01.dta.**
- Suponiendo que el archivo **pei_08.dta** no está en el actual directorio. Para hacer que cambie el uso de su directorio de:

  use “/Users/DCA/Desktop/eval/data/pei_01.dta”

  ahora a **Desktop**, teclee en *ventana de comando* **STATA:**

  use “/Users/DCA/Desktop/pei_01.dta”

```
. use "/Users/DCA/Desktop/eval/data/pei_01.dta"

.
Command
```

Con vista de la ventana de variables de **STATA.**

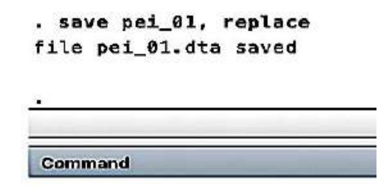
Variables	
Name	Label
id	ID
year	year of obser...
region	village ID
cluster	idcluster
ageceo	age of ceo
genceo	gender of ce...
educeo	education of...
smesize	sme size bas...
smefac	sme assets b...
incinnov	sme incinnov
expopn	sme opn exp...
expinnov	sme innovatio...
exptot	total expendit...
dmmfd	innovation ma...
dfmfd	innovation fe...
weight	innovation sa...
bdbaccess	sme broad ba...
pcirr	Proportion of...
raw1	raw1
raw2	raw2
raw3	raw3
raw4	raw4
raw5	raw5
raw6	raw6

Ejemplo 2. Guardando el conjunto de datos (*Saving data set*)

Se cuenta con las siguientes alternativas:

- Si realiza cambios en un archivo de datos **STATA** abierto y desea guardar esos cambios, puede hacerlo, usando el comando *save* de **STATA**. Por ejemplo, el siguiente comando guarda el archivo **pei_01.dta**:

  . save pei_01, replace



- La opción *replace* le indica a **STATA** que sobrescriba la versión el original con la nueva versión. Si no quieres perder la versión original, tienes para especificar un nombre de archivo diferente en el comando *save*.

Ejemplo 3. Saliendo de STATA (*exiting STATA*)

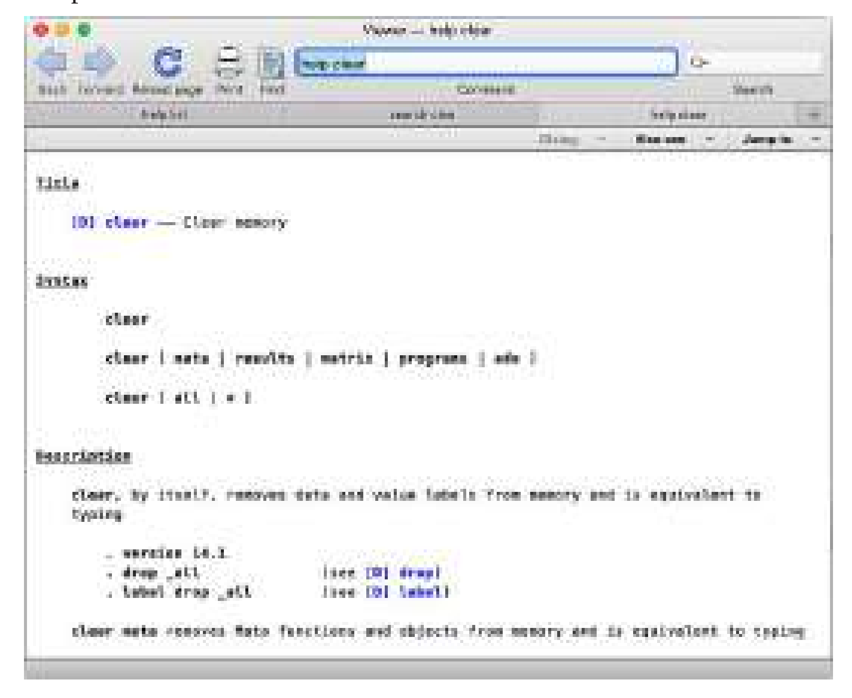
Una manera fácil de salir de **STATA** es emitir el comando *exit*. Sin embargo, si tiene un conjunto de datos sin guardar abierto, **STATA** emitirá el siguiente mensaje de error:



Ejemplo 4. Apoyo de STATA (*Help*)

STATA viene con un excelente, enorme y útil conjunto de manuales multivolumen. Con acceso a la Web, incluso dispone de un conjunto más grande de macros y otra información útil. Desde **STATA**, puede escribir el comando *help* y la palabra clave para solicitar información de ayuda sobre. Este comando, solo funciona si escribe el comando el nombre completo nombre o palabra clave sin abreviaturas. Por ejemplo:

. help clear



. help clea



En **STATA**, por cierto, un valor faltante se representa con un punto (.). Se considera un valor faltante más grande que cualquier número. El comando *summarize* ignora las observaciones con valores faltantes, y el comando *tabulate* hace lo mismo, a menos que se vea obligado a incluir los valores faltantes.

Ejemplo 5. Apoyo de STATA (*Lookup/Search*)

Si no puede recordar el nombre completo del comando o la palabra clave, o si no está seguro acerca de qué comando desea, puede usar el comando *lookup* o *search* seguido el nombre del comando o palabra clave, escribiéndose así:

. search mem



Este comando, enumerará todos los asociados con esta palabra clave y mostrará una breve descripción de cada uno de ellos. Entonces puede elegir el comando que cree que es relevante y utilizar la ayuda para obtener la referencia específica. El sitio web https://www.stata.com/ tiene excelentes facilidades de ayuda, como un tutorial en línea y preguntas frecuentes **(FAQ).**

Ejemplo 6. Notas sobre los comandos STATA

Aquí hay algunos comentarios generales sobre los comandos **STATA**:

- Los comandos **STATA** se escriben en *minúsculas.*
- Todos los nombres, incluidos los comandos o nombres de variables, *se pueden abreviar siempre que no existe ambigüedad*. Por ejemplo: *describe*, y simplemente *hacen el mismo trabajo porque no existe confusión.*
- Además de escribir, se pueden usar algunas teclas para representar algunos comandos **STATA** o secuencias. Los más importantes de ellos son las teclas *Page-Up* y *Page-Down*. Para mostrar el comando anterior en la *ventana de comando* **STATA** puede presionar la tecla *Page-Up*. Puedes seguir haciéndolo, hasta que el primer comando de la sesión aparece. Del mismo modo, la tecla *Page-Down* muestra el comando que sigue al que actualmente usa, en la misma *ventana de comandos* **STATA.**
- Al oprimir una vez en un comando, en la ventana de revisión (*review window*), lo colocará también, en el *ventana de comando* **STATA**. Al hacer oprimir doble, le indicará a **STATA** que ejecute el comando. Esto puede ser útil, cuando los comandos deben repetirse o editarse ligeramente en la *ventana de comando* **STATA.**

Ejemplo 7. Listando las variables (*Listing variables*)

Para ver todas las variables en el conjunto de datos, utilice el comando *describe* (completo o abreviado). Este comando proporciona información sobre el conjunto de datos y *enumera* todas las variables. Para ver solo una variable o lista de variables, utilice el comando de *describe* seguido del nombre o los nombres de las variables:

```
. desc id region
```

```
. desc id year region

              storage  display    value
variable name   type   format     label      variable label

id              long   %12.0g                HH ID
year            byte   %8.0g                 Year of observation
region          byte   %8.0g                 Village ID

.
Command
```

Como puede ver, el comando *describe* también muestra el *tipo y la longitud* de la variable, como una breve descripción de la variable (si está disponible). Deben tenerse en cuenta los siguientes puntos:

- *Puede abreviar una lista de variables* escribiendo el nombre tan solo de la primera y la última variable, separados por un guión (-); la ventana de variables (*variable window*) muestra el orden en que las variables se almacenan. Por ejemplo, para ver todas las variables desde id hasta raw se escribe:

. desc id-raw6

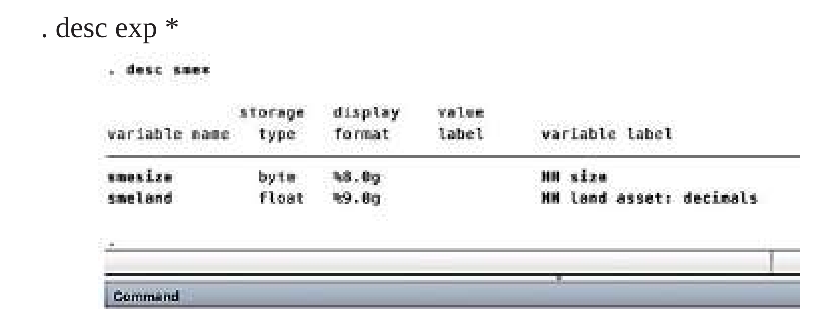
```
. desc id-raw6

              storage  display    value
variable name   type   format     label      variable label

id              long   %12.0g                HH ID
year            byte   %8.0g                 Year of observation
region          byte   %8.0g                 Village ID
cluster         byte   %8.0g                 idcluster
agehead         byte   %8.0g                 Age of HH head: years
genhead         byte   %8.0g                 Gender of HH head: 1=M, 0=F
eduhead         byte   %8.0g                 Education of HH head: years
[illegible]     byte   %8.0g                 HH size
hhland          float  %9.0g                 HH land asset: decimals
[illegible]     double %12.0g                HH total asset: Tk.
[illegible]     float  %9.0g                 HH per capita food expenditure: Tk/year
[illegible]     float  %9.0g                 HH per capita nonfood expenditure: Tk/year
[illegible]     float  %9.0g                 HH per capita total expenditure: Tk/year
dmmfd           byte   %8.0g                 HH has male microcredit participant: 1=Y, 0=N
dfmfd           byte   %8.0g                 HH has female microcredit participant: 1=Y, 0=N
weight          float  %9.0g                 HH sampling weight
—more—
Command
```

- El símbolo de comodín (*) es útil para guardar algo de escritura. Por ejemplo, para ver todas las variables que comienzan con sme, se escribe:

. desc exp *



```
. desc sme*

              storage  display    value
variable name   type   format     label      variable label

smesize         byte   %8.0g                 HH size
smeland         float  %9.0g                 HH land asset: decimals

.
Command
```

- Puede abreviar una variable o lista de variables de esta manera, en cualquier comando (donde tenga sentido), no solo con *describe.*

Ejemplo 8. Listando los datos (*Listing data*)

Para ver los datos al momento, almacenados en las variables, utilice el comando *list* (abreviado como *l*). Si escribe el comando *list* por sí mismo, **STATA** mostrará valores para todas las variables y todas las observaciones, que pueden no ser deseables para ningún propósito práctico (así, se usa la combinación *Ctrl-Break*, para evitar que los datos se desplacen sin fin por la pantalla). Por lo general, se desea ver los datos de ciertas variables y de ciertas observaciones. Esto se logra escribiendo un comando *list* con una lista de variables y con condiciones. El siguiente comando enumera todas las variables de las primeras tres observaciones:

. list in 1/3

```
. list in 1/3

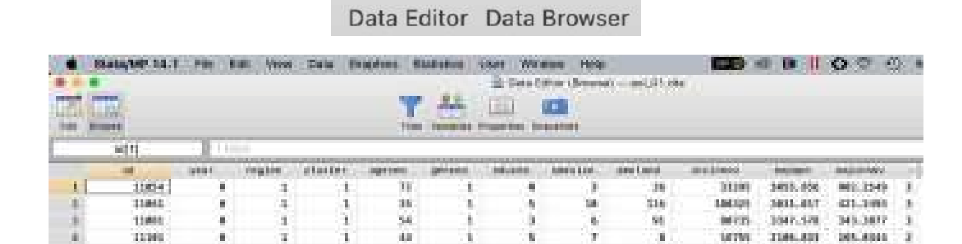
     +------------------------------------------------------------------------------------+
  1. |    id | year | region | cluster | ageceo | genceo | educeo | smesize | smeland |
     | 11054 |    0 |      1 |       1 |     72 |      1 |      0 |       3 |      36 |
     |------------------------------------------------------------------------------------|
     | incinnov |   expopn | expinnov |   exptot | dmmfd | dfmfd |   weight | vaccess | pcirr |
     |    33295 | 3055.856 | 902.2549 | 3958.112 |     0 |     0 | 1.042735 |       1 |   .45 |
     |------------------------------------------------------------------------------------|
     |     raw1     |     raw2     |     raw3     |     raw4     |     raw5     |     raw6     |
     |       11     |      8.5     |       12     |        8     |      2.5     |       55     |
     +------------------------------------------------------------------------------------+

     +------------------------------------------------------------------------------------+
  2. |    id | year | region | cluster | ageceo | genceo | educeo | smesize | smeland |
     | 11061 |    0 |      1 |       1 |     35 |      1 |      5 |      10 |     116 |
     |------------------------------------------------------------------------------------|
     | incinnov |   expopn | expinnov |   exptot | dmmfd | dfmfd |   weight | vaccess | pcirr |
     |   180325 | 3031.017 | 421.2493 | 3452.266 |     1 |     0 | 1.011396 |       1 |   .45 |
     |------------------------------------------------------------------------------------|
     |     raw1     |     raw2     |     raw3     |     raw4     |     raw5     |     raw6     |

Command
```

Aquí, **STATA** muestra todos los registros que comienzan con la observación **1** y terminan con observación **3**.

STATA, también puede mostrar los datos como una *hoja de cálculo*. Para hacerlo, use los dos íconos en la *barra de herramientas* llamada *Data Editor* y *Data Browser*). Oprimiendo una vez, hará que aparezca una nueva ventana donde los datos se mostrarán como una tabla, con los registros como filas y las variables como columnas. *Data Browser* solo mostrará los datos, mientras que para editarlos necesitará de *Data Editor*. Los comandos *Edit y Browser,* también abrirán la ventana de la *hoja de cálculo*.



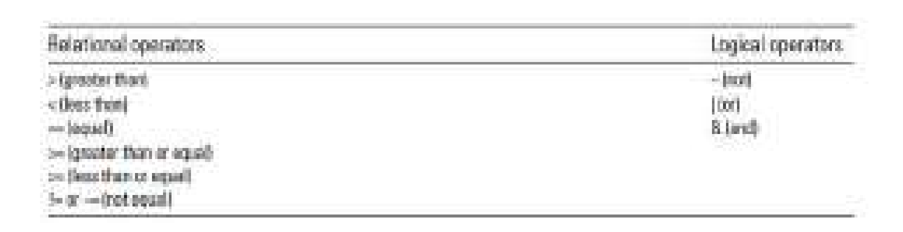


Ejemplo 9. Listando datos específicos

El siguiente comando enumera el tamaño de una pyme de base tecnológica y la educación del ceo de género femenino menor de 45 años:

```
. list smesize educeo if (genceo==0 & ageceo<45)
```

La declaración anterior usa dos *operadores relacionales* (== y <) y un operador lógico (&). Los *operadores relacionales* imponen una condición en una variable, mientras que los *operadores lógicos* combinar dos o más operadores relacionales. Abajo se muestra la relación y operadores lógicos utilizados en **STATA.**



<table>
<tr><th>Relational operators</th><th>Logical operators</th></tr>
<tr><td>> (greater than)</td><td>~ (not)</td></tr>
<tr><td>< (less than)</td><td>| (or)</td></tr>
<tr><td>== (equal)</td><td>& (and)</td></tr>
<tr><td>>= (greater than or equal)</td><td></td></tr>
<tr><td><= (less than or equal)</td><td></td></tr>
<tr><td>!= or ~= (not equal)</td><td></td></tr>
</table>

Puede usar operadores relacionales y lógicos en cualquier comando Stata (donde haga sentido), no solo en el comando *list*.

```
. list smesize educeo if (genceo==0 & ageceo<45)

     +------------------+
     | smesize   educeo |
     |------------------|
 35. |       4        0 |
 44. |       7        2 |
 54. |       5        0 |
 68. |       2        0 |
 99. |       6        0 |
     |------------------|
108. |       2        0 |
135. |       4        0 |
170. |       2        0 |
180. |       6        0 |
228. |       3        0 |
     |------------------|
229. |       3        0 |
234. |       5        0 |
274. |       5        8 |
389. |       3        0 |

Command
```

Ejemplo 10. Sumando datos (*Summarizing data*)

El comando *summarize* (el cual puede abreviarse como *sum*) calcula y muestra algunas sumas de estadísticas, que incluyen *medias y desviaciones estándar.* Sino se especifica alguna variable, este comando aplicará para todas para todas las variables incluídas en el conjunto de datos.

Como ejemplo, se muestra la suma de las variables tamaño de la pyme de base tecnológica y la educación del ceo, como sigue:

. sum smesize educeo

```
. sum smesize educeo [aw=weight]

    Variable |  Obs      Weight        Mean   Std. Dev.       Min        Max
-------------+-------------------------------------------------------------
     smesize |  826   822.32837    5.728505   2.567759          1         19
      educeo |  826   822.32837    2.657877   3.615297          0         16

.

Command
```

STATA excluye cualquier observación que tenga un valor faltante para las variables que se *suman* a partir de este cálculo (los valores faltantes se analizan más adelante). Si se requieres saber la mediana y los percentiles de una variable, agregue la opción *detail* (abreviado *d*), por ejemplo:

. sum smesize educeo, d

```
. sum smesize educeo, d

                          HH size
-------------------------------------------------------------
      Percentiles      Smallest
 1%            2              1
 5%            2              1
10%            3              1       Obs                 826
25%            4              1       Sum of Wgt.         826

50%            5                      Mean           5.470944
                        Largest       Std. Dev.      2.378643
75%            7             15
90%            8             17       Variance       5.657943
95%           10             17       Skewness       1.308173
99%           14             19       Kurtosis       6.355481

                 Education of HH head: years
-------------------------------------------------------------
—more—
Command
```

Ejemplo 11. Ponderaciones (*Weight*)

Una gran fortaleza de **STATA** es que permite el uso de ponderaciones (*weights*). La opción de *weight* es útil, *si la probabilidad de muestreo de una observación es diferente de la de otra.* Por ejemplo, en nuestro caso es de suponer que en las pymes de base tecnológica (**pbt**), el marco de muestreo estaría estratificado, donde el primer muestreo primario las unidades (a menudo regiones) se muestrean y dependen de la selección de

las unidades primarias, de las que se extraen unidades de muestreo secundarias (a menudo **pbt** de cierto tamaño). Las encuestas, generalmente deben aplicar ponderaciones para corregir las diferencias de diseño de muestreo y, a veces problemas de recolección de datos. La implementación en **STATA** es sencilla:

. sum smesize educeo [aw=weight]

```
. sum smesize educeo

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
     smesize |       826    5.470944    2.378643          1         19
      educeo |       826    2.405569    3.333702          0         16

.
Command
```

Aquí, la variable *weight* tiene la información sobre el peso que se debe dar a cada observación y *aw* es una opción de **STATA** para incorporar el peso en el cálculo. El uso de ponderaciones se analiza más adelante en ejercicios de capítulos posteriores. Para las variables que son cadenas (*strings*), el comando *summarize* no aplica para producir datos de estadística descriptiva, excepto que el número de observaciones es cero. También para *variables categóricas* (por ejemplo, educación secundaria = 1, educación superior = 2, etc.), interpretar el resultado del comando *summarize* puede ser dificil. En ambos casos, una tabulación completa puede ser más significativa, lo cual se discute enseguida.

En la mayoría del levantamiento de encuestas, las observaciones se seleccionan mediante un proceso aleatorio y puede tener diferentes probabilidades de selección. Por lo tanto, se deben usar *ponderaciones* que son igual a la inversa de la probabilidad de ser muestreada. Una ponderación de *wn* para el *enésima* observación significa, en términos generales, que la observación *n* representa elementos *wj* en la población de la cual se extrajo la muestra. Omitir las ponderaciones de muestreo en el análisis, general-

mente proporciona estimaciones sesgadas, que pueden estar lejos de los valores verdaderos.

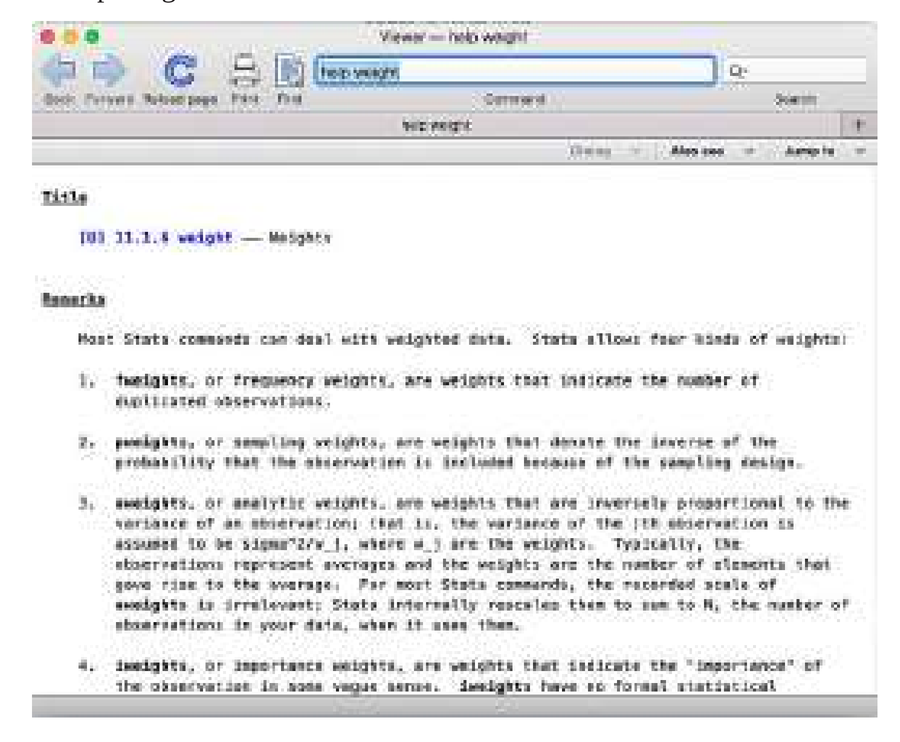
Por lo general, son necesarios varios ajustes a las ponderaciones, son necesarios posteriores al muestreo. Por ejemplo, la ponderación de muestreo de las pymes de base tecnológica (**pbt**) de la **pei_01.dta** es la ponderación correcta a usar, cuando se utiliza para sumarizar los datos relacionados.

STATA tiene cuatro tipos de ponderaciones:

- *Ponderación por frecuencia* (frequency weights. *fweight*), que indica cuántas observaciones en la población están representados por cada una de la muestra y deben tomar valores enteros.
- *Ponderación analítica* (analytic weights. *aweight*) son especialmente apropiados, cuando se trabaja con datos que contienen promedios (por ejemplo, ingreso promedio per cápita en una pyme de base tecnológica). La variable de ponderación, es proporcional al número de pymes sobre las cuales se calculó el promedio (por ejemplo, número de miembros de una pyme). Técnicamente, *la ponderación analítica están en proporción inversa a la varianza de una observación* (es decir, una mayor ponderación significa que la observación se basó en más información y, por lo tanto, es más confiable en el sentido de tener menos varianza).
- Las *ponderaciones de muestreo* (sampling weights. *pweight*) son lo inverso de la probabilidad de selección, debido al diseño de la muestra.
- Las ponderaciones de importancia (importance weights. *iweight*) que indican la importancia relativa de la observación.

Los más comúnmente usados son pweight y aweight. Más información sobre las ponderaciones, se pueden obtener escribiendo:

. help weight



Los siguientes comandos muestran más sobre la aplicación de ponderaciones:

. tabstat smesize [aweight=weight], statistics(mean sd) by(dfmfd)
. table dfmfd [aweight=weight], contents(mean smesize sd smesize)

```
. tabstat smesize [aweight=weight], statistics(mean sd) by(dfmfd)

Summary for variables: smesize
     by categories of: dfmfd (innovation female microcredit participant: 1=Y, 0=N)

   dfmfd |      mean        sd
---------+--------------------
       0 |  5.816999  2.672921
       1 |  5.358539  2.036723
---------+--------------------
   Total |  5.728505  2.567759
------------------------------

. table dfmfd [aweight=weight], contents(mean smesize sd smesize)

----------------------------------------
innovatio |
n female  |
microcred |
it        |
participa |
nt: 1=Y,  |
0=N       |  mean(smesize)    sd(smesize)
----------+-----------------------------
        0 |          5.817       2.672921
        1 |        5.35854       2.036723
----------------------------------------

.
```

Command

Ejemplo 12. Ordenamiento con ponderación (*Sort / Weight*)

A menudo, se requieren ver estadísticas sumarizadas por grupos de ciertas variables, no solo para todo el conjunto de datos. Supongamos que quiere ver el tamaño de la pyme de base tecnológica y la educación del CEO para *participantes y no participantes*. Así:

- Primero, ordene los datos por variable de grupo (en este caso, *dfmfd)*.
- Puede verificar el ordenamiento con el comando *describe,* después de abrir cada archivo.
- El comando *describe*, después de enumerar todas las variables, indica si el conjunto de datos está ordenado por alguna variable.
- Si no hay información del ordenamiento, o el conjunto de datos se ordena por una variable que es diferente de la que desea, puede

utilizar el comando *sort* y posteriormente, guardar el conjunto de datos en este formulario.

- Los siguientes comandos ordenan el conjunto de datos por la variable *dfmfd* y muestran la sumarización estadística del tamaño de la yme con base tecnológica y la educación CEO para *participantes y no participantes:*

. sort dfmfd
. by dfmfd: sum smesize educeo [aw=weight]

```
. sort dfmfd

. by dfmfd: sum smesize educeo [aw=weight]

-> dfmfd = 0

    Variable |     Obs      Weight        Mean   Std. Dev.       Min        Max
-------------+------------------------------------------------------------------
     smesize |     539  663.598263    5.816999   2.672921          1         19
      educeo |     539  663.598263     2.85233   3.771715          0         16

-> dfmfd = 1

    Variable |     Obs      Weight        Mean   Std. Dev.       Min        Max
-------------+------------------------------------------------------------------
     smesize |     287  158.730107    5.358539   2.036723          1         14
      educeo |     287  158.730107     1.84493   2.734086          0         12

Command
```

Ejemplo 13. Uso del comando tabstat

Una alternativa útil al comando *summary* es el comando *tabstat*, que le permite especificar la lista de estadísticas que desea mostrar en una sola tabla. Puede estar condicionado por otra variable. El siguiente comando muestra la media y desviación estándar del tamaño de pyme de base tecnológica y educación del CEO por la variable dfmfd:

. tabstat smesize educeo, statistics(mean sd) by(dfmfd)

```
. tabstat smesize educeo, statistics(mean sd) by(dfmfd)

Summary statistics: mean, sd
  by categories of: dfmfd (HH has female microcredit participant: 1=Y, 0=N)

   dfmfd |   smesize    educeo
---------+--------------------
       0 |  5.560297  2.669759
         |  2.582033  3.559692
---------+--------------------
       1 |  5.303136  1.909406
         |  1.932992  2.808877
---------+--------------------
   Total |  5.470944  2.405569
         |  2.370643  3.333702
------------------------------

.
Command
```

Ejemplo 14. Distribuciones de frecuencia (*Tabulations*)

Las distribuciones de frecuencia y las tabulaciones cruzadas (*cross-tabulations*) a menudo son necesarias. El comando *tabulate* (abreviado *tab*) se realiza, como sigue:

. tab dfmfd

```
. tab dfmfd

  innovation
      female
 microcredit
 participant
 : 1=Y, 0=N |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |        539       65.25       65.25
          1 |        287       34.75      100.00
------------+-----------------------------------
      Total |        826      100.00

.
Command
```

El siguiente comando, proporciona la distribución de género de los CEO de las pymes de base tecnológica (**pbt**) participantes:

. tab genceo if dfmfd==1

```
. tab genceo if dfmfd==1

  gender of |
ceo: 1=M,   |
       0=F  |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |         23        8.01        8.01
          1 |        264       91.99      100.00
------------+-----------------------------------
      Total |        287      100.00

.
Command
```

Tenga en cuenta aquí, el uso del signo = = . Esto indica que si la variable es igual a uno, entonces se realiza la tabulación. El comando *tabulate* también puede ser utilizado para mostrar una distribución bidireccional (*two-way distribution*). Por ejemplo, si desea verificar si existe algún *sesgo* de género en la educación de los CEO de las pymes de base tecnológica (**pbt**), tiene el comando:

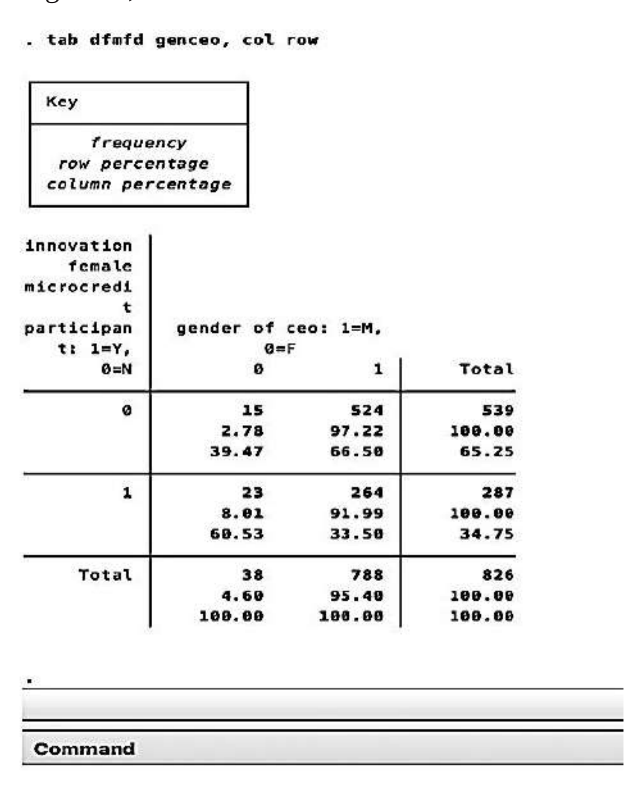
. tab educeo genceo

```
. tab educeo genceo

 education |  gender of ceo: 1=M,
   of ceo: |         0=F
     years |         0          1 |     Total
-----------+----------------------+----------
         0 |        32        414 |       446
         1 |         0         27 |        27
         2 |         1         56 |        57
         3 |         2         39 |        41
         4 |         1         43 |        44
         5 |         1         77 |        78
         6 |         0         23 |        23
         7 |         0         21 |        21
         8 |         1         26 |        27
         9 |         0         15 |        15
        10 |         0         26 |        26
        11 |         0          2 |         2
        12 |         0         14 |        14
--more--
Command
```

Para ver porcentajes por fila o columnas, agregue opciones al comando *tabulate*:

. tab dfmfd genceo, col row



```
. tab dfmfd genceo, col row

Key
frequency
row percentage
column percentage
```

innovation female microcredit participant: 1=Y, 0=N	gender of ceo: 1=M, 0=F 0	1	Total
0	15 2.78 39.47	524 97.22 66.50	539 100.00 65.25
1	23 8.01 60.53	264 91.99 33.50	287 100.00 34.75
Total	38 4.60 100.00	788 95.40 100.00	826 100.00 100.00

.

Command

Ejemplo 15. Estadística descriptiva (*Table command*)

Otro comando muy conveniente es *table*, que combina características de los comandos *sum* y *tab*. Además, muestra los resultados en una forma más presentable. El siguiente comando *table*, muestra la *media* del tamaño de la pyme de base tecnológica y de la educación del CEO, por su participación en programa de microfinanzas:

.table dfmfd, c (mean smesize mean educeo)

```
. table dfmfd, c (mean smesize mean educeo)

innovatio
n female
microcred
it
participa
nt: 1=Y,
0=N          mean(smesize)   mean(educeo)

        0           5.5603              3
        1          5.30314              2

.
Command
```

Ejemplo 16. Despliegue de punto decimal

Los resultados son los esperados. Pero, ¿por qué la media de educeo se muestra como entero y no una fracción? Esto ocurre porque la variable educeo se almacena como un número entero, y **STATA** simplemente trunca los números después del decimal. Observe la descripción de la variable:

.d educeo

```
. d educeo

              storage  display    value
variable name   type   format     label      variable label

educeo          byte   %2.0f                 education of ceo: years

.
Command
```

Tenga en cuenta que educeo, es una variable con punto flotante: su formato (% 2.0f) muestra que sus dígitos ocupan dos lugares y no tiene ningún dígito después del decimal. Es posible forzar cambiar dicho formato; suponga que desea que muestre dos lugares después del decimal para una visualización de tres dígitos, el cual se realiza como sigue:

. format educeo %3.2f
. table dfmfd, c (mean smesize mean educeo)

```
. format educeo %3.2f

. table dfmfd, c (mean smesize mean educeo)
```

innovation female microcredit participant: 1=Y, 0=N	mean(smesize)	mean(educeo)
0	5.5603	2.67
1	5.30314	1.91

```
.
.
```

Command

El formato solo cambia la visualización de la variable, no su representación interna de la variable en la memoria. El comando *table*, puede mostrar hasta cinco estadísticas de las variables como de la media, la suma, el mínimo o máximo. Se pueden mostrar tablas bidimensionales, tridimensionales o incluso de dimensiones superiores. Por ejemplo, mostramos una tabla bidimensional que desglosa la educación del CEO, no solo por región, sino también género del mismo.

. table dfmfd genceo, c (mean smesize mean educeo)

```
. table dfmfd genceo, c (mean smesize mean educeo)

----------------------------------
innovatio  |
n female   |
microcred  |
it         |
participa  |   gender of ceo:
nt: 1=Y,   |      1=M, 0=F
0=N        |        0        1
-----------+----------------------
         0 |  3.13333  5.62977
           |     0.53     2.73
           |
         1 |   4.6087  5.36364
           |     0.74     2.01
----------------------------------

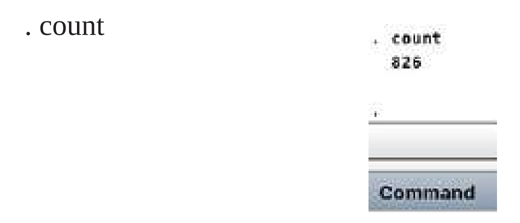
.

Command
```

Ejemplo 17. Comando contar (*Count*)

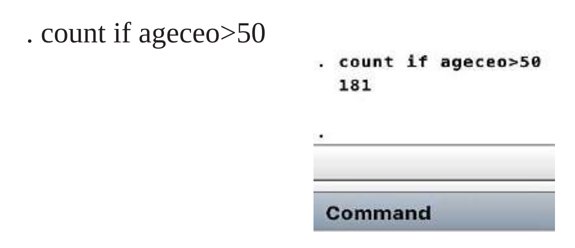
El comando *count*, se utiliza para contar el número de observaciones en el conjunto de datos:

. count



El comando *count*, se puede usar con condiciones. Por ejemplo, el siguiente comando proporciona el número de pymes de base tecnológica (**pbt**) cuyos CEO tienen una edad de más de más de 50 años:

. count if ageceo>50



Hasta ahora, la discusión se ha limitado a los comandos **STATA** que muestran información en el datos de diferentes maneras *sin cambiar los datos*. En realidad, las sesiones más frecuentes de su uso, implica hacer cambios en los datos (por ejemplo, crear nuevas variables o cambiar valores de variables existentes). Los siguientes ejercicios demuestran cómo esos cambios se puede incorporar en **STATA.**

Ejemplo 18. Generando nuevas variables

En **STATA**, el comando *genérate* (abreviado *gen*) crea nuevas variables, mientras que el comando *replace* cambia los valores de una variable existente. Los siguientes comandos crean una nueva variable llamada oldceo y posteriormente, establecen su valor a una de las pymes con CEO de edad de más 50 años y a cero de lo contrario:

. gen oldceo=1 if ageceo>50

```
. gen oldceo=1 if ageceo>50
(645 missing values generated)
```

Command

oldceo
1
.
1
.
.
.
1

. replace oldceo=0 if ageceo<=50



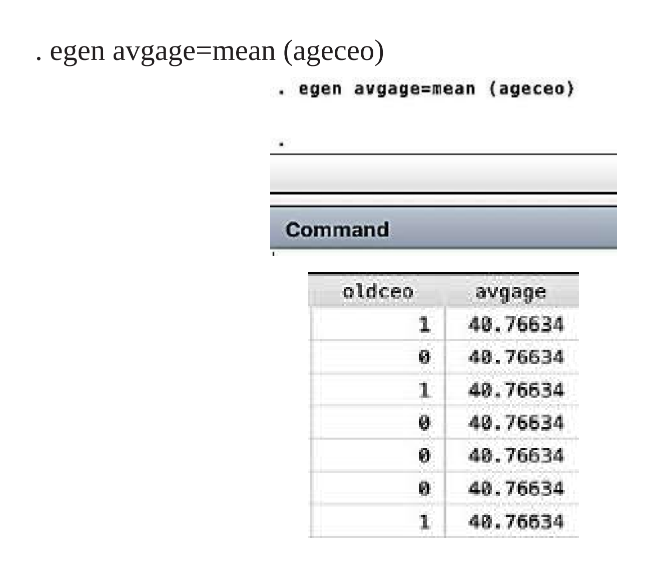
Así, para cada observación, el comando *gen* verifica la condición (por ejemplo, si el CEO de la pyme de base tecnológica, tiene más de 50 años) y establece el valor de la variable *oldceo*, uno para esa observación, si la condición es verdadera o a valor perdido, de lo contrario. El comando *replace* funciona de manera similar. Después del comando *generate*, **STATA** indica que 645 observaciones *no cumplieron con la condición*, y después del comando *replace* **STATA** indica que esas 645 observaciones tienen nuevos valores (cero en este caso). Vale la pena señalar los siguientes puntos:

- Si se emite un comando *gen* o *replace* sin ninguna condición, ese comando se aplica a todo el archivo de datos.
- Al usar el comando *gen*, se debe tener cuidado de manejar los valores faltantes correctamente.
- El lado derecho del signo = en los comandos *gen* o *replace* puede ser cualquier expresión que involucre nombres de variables, no solo un valor. Así, por ejemplo, el comando *gen young = (ageceo <= 32)* crearía una variable llamada *young* que tomaría el valor de los CEO de pyme tecnológica que tengan 32 años o menos de edad (si la

expresión entre paréntesis, es verdadera) y un valor de cero en caso contrario.

- El comando replace, se puede usar para cambiar los valores de cualquier variable existente, independientemente del comando *gen*.
- Una extensión del comando *gen* es *egen*. Al igual que el comando *gen*, el comando *egen* puede crear variables para almacenar estadísticas descriptivas, como: *la media*, *suma*, *máximo y mínimo*. La característica más poderosa del comando *egen* es su capacidad para crear estadísticas que involucran *múltiples observaciones*. Por ejemplo, el siguiente comando crea una variable *avgage* que contiene la edad promedio del CEO de cada pyme de base tecnológica, del archivo **pei_01.dta**:

. egen avgage=mean (ageceo)



```
. egen avgage=mean (ageceo)

.
```

Command

oldceo	avgage
1	40.76634
0	40.76634
1	40.76634
0	40.76634
0	40.76634
0	40.76634
1	40.76634

Todas las observaciones en el conjunto de datos, obtienen el mismo valor para *avgage*. El siguiente comando crea las mismas estadísticas, pero esta vez para pymes de base tecnológica (**pbt**) de hombres y mujeres por separado:

. egen avgagemf=mean (ageceo), by (genceo)

```
. egen avgagemf=mean (ageceo), by (genceo)

.
Command
```

oldceo	avgage	avgagemf
1	40.76634	40.70304
0	40.76634	40.70304
1	40.76634	40.70304
0	40.76634	40.70304
0	40.76634	40.70304
0	40.76634	40.70304
1	40.76634	40.70304

Ejemplo 19. Etiquetando variables (*Labeling variables*)

Puede etiquetar a las variables para darles una descripción. Por ejemplo, la variable *oldceo* no tiene ninguna etiqueta ahora. Puede adjuntar una etiqueta a esta variable escribiendo:

. label var oldceo “CEO is over 50: 1=Y, 0=N”



Ahora para ver lo nueva etiqueta, escriba lo siguiente:

.des oldceo

```
. des oldceo

               storage   display    value
variable name    type    format     label      variable label

oldceo           float   %9.0g                 CEO is over 50: 1=Y, 0=N

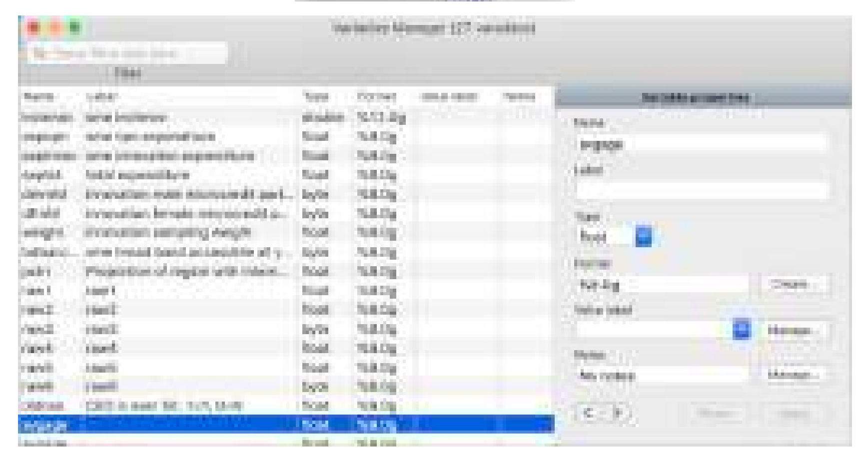
.
Command
```

Ejemplo 20. Etiquetando variables (*Labeling variables*) desde barra de herramientas

También, es posible corregir desde la barra de herramientas oprimiendo:

Data→Variables Manager:



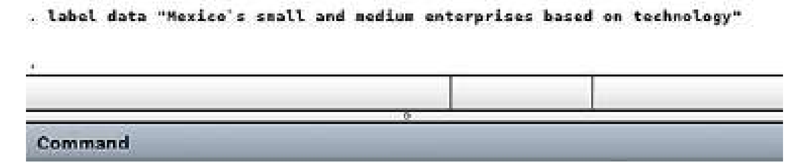


Hacer las modificaciones necesarias y oprimir el botón Apply para finalizar.

Ejemplo 21. Etiquetando datos (*Labeling data*)

Se pueden crear otros tipos de etiquetas. Para adjuntar una etiqueta a todo el conjunto de datos, teclee:

. label data “Mexico`s small and medium enterprises based on technology”



Para ver la etiqueta, teclee:

. des

```
. des

Contains data from /Users/DCA/Desktop/eval/data/pei_01.dta
  obs:           826                          Mexico`s small and medium
                                                enterprises based on
                                                technology
 vars:            27                          14 Jun 2020 19:00
 size:        62,776
--------------------------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
--------------------------------------------------------------------------------
id              long    %12.0g                ID
year            byte    %8.0g                 year of observation
region          byte    %8.0g                 village ID
cluster         byte    %8.0g                 idcluster
ageceo          byte    %8.0g                 age of ceo
genceo          byte    %8.0g                 gender of ceo: 1=M, 0=F
educeo          byte    %8.0g                 education of ceo: years
smesize         byte    %8.0g                 sme size based on personel
smefac          float   %9.0g                 sme assets based on facilities
incinnov        double  %12.0g                sme incinnov
expopn          float   %9.0g                 sme opn expenditure
expinnov        float   %9.0g                 sme innovation expenditure
exptot          float   %9.0g                 total expenditure
dmmfd           byte    %8.0g                 innovation male microcredit
                                                participant: 1=Y, 0=N
dfmfd           byte    %8.0g                 innovation female microcredit
                                                participant: 1=Y, 0=N
weight          float   %9.0g                 innovation sampling weight
--more--

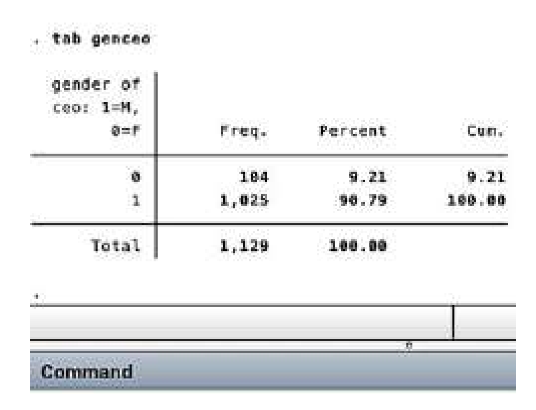
Command
```

Observe que el conjunto de datos de **pei_01.dta** tiene el nombre de “Mexico`s small and medium enterprises based on technology”, el cual permite asignar nombres al conjunto de datos.

Ejemplo 22. Etiquetando valores de variables (*Labeling values of variables*)

Las variables categóricas, como *genceo* (1 = hombre, 0 = mujer), pueden tener etiquetas que ayudan a recordar cuáles son las categorías. Por ejemplo, usando **pei_08.dta**, tabular la variable *genceo* muestra solo valores 0 y 1:

.tab genceo

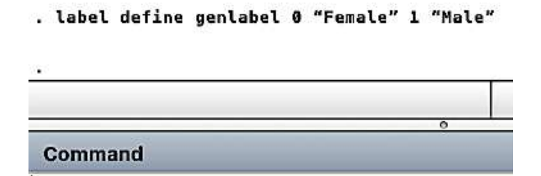

. tab genceo

gender of ceo: 1=M, 0=F	Freq.	Percent	Cum.
0	104	9.21	9.21
1	1,025	90.79	100.00
Total	1,129	100.00	

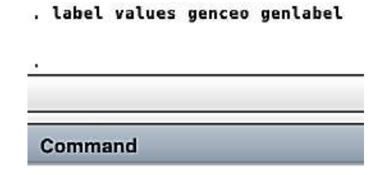
Para adjuntar etiquetas a los valores de una variable, y volviendo a **pei_01.dat,** se deben hacer dos cosas:

- Primero, definir una etiqueta de valor.
- Posteriormente, asigne esta etiqueta a la variable. Usando las nuevas categorías para *genceo*, teclee:

. label define genlabel 0 “Female” 1 “Male”

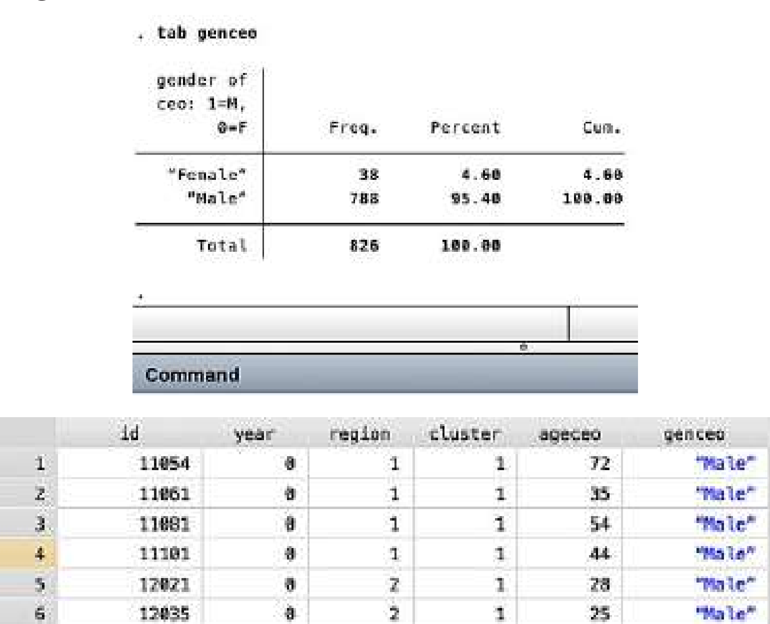


. label values genceo genlabel



Ahora, para ver las etiquetas, escriba:

. tab genceo



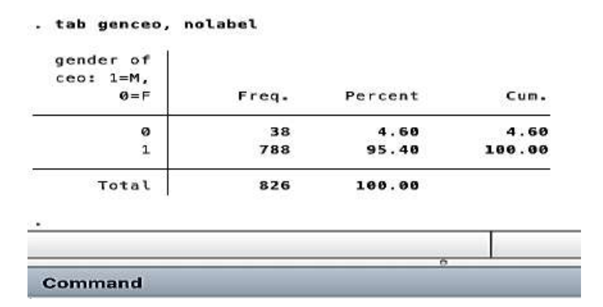
. tab genceo

gender of ceo: 1=M, 0=F	Freq.	Percent	Cum.
"Female"	38	4.60	4.60
"Male"	788	95.40	100.00
Total	826	100.00	

	id	year	region	cluster	ageceo	genceo
1	11054	0	1	1	72	"Male"
2	11061	0	1	1	35	"Male"
3	11081	0	1	1	54	"Male"
4	11101	0	1	1	44	"Male"
5	12021	0	2	1	28	"Male"
6	12035	0	2	1	25	"Male"

Si desea ver los valores reales de la variable genceo, que siguen siendo ceros y unos, puede agregar una opción para no mostrar las etiquetas asignadas a los valores de variable. Por ejemplo, teclee:

. tab genceo, nolabel



. tab genceo, nolabel

gender of ceo: 1=M, 0=F	Freq.	Percent	Cum.
0	38	4.60	4.60
1	788	95.40	100.00
Total	826	100.00	

Ejemplo 23. Soltar variables así como observaciones

Las variables y observaciones de un conjunto de datos, se pueden seleccionar mediante el uso de los comandos *keep* o *drop.* Suponga que tiene de la base de datos .dta seis variables: var1, var2,…, var6 y desea mantener un archivo con solo tres de ellos (por ejemplo, var1, var2 y var3). Así, se es posible usar cualquiera de los siguientes dos comandos:

- "keep var1 var2 var3" (or "keep var1-var3" si es este el orden deseado).
- "drop var4 var5 var6" (or "drop var4-var6" si es edste e orde deseado).

Tenga en cuenta el uso de un guión (-) en ambos comandos. Es una buena práctica, usar el comando que implica el menor tecleo posible de las variables (y, por lo tanto, tener menos riesgo de error). También es posible usar use operadores relacionales o lógicos. Por ejemplo, el siguiente comando muestra todos los ceo de pymes de base tecnológica (**pbt**) con edad de 80 años o más:

drop if ageceo>=80

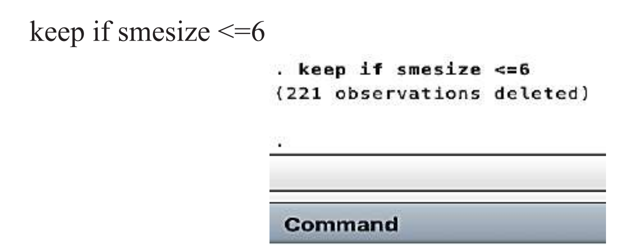
```
. drop if ageceo>=80
(0 observations deleted)

.

Command
```

Y este comando, por ejemplo, mantiene esas observaciones donde el tamaño de la pyme de base tecnológica, es de seis o menos miembros:

keep if smesize <=6



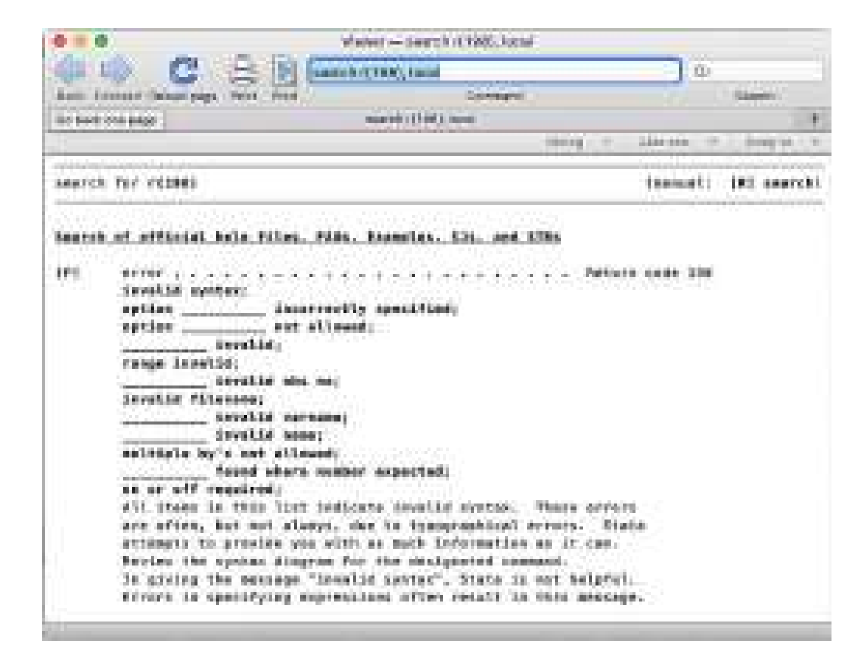
Los dos comandos anteriores eliminan o mantienen todas las variables según las condiciones. Así, no es posible incluir una lista de variables en un comando *drop* o *keep* que también utiliza condiciones. Por ejemplo, el siguiente comando provocará errores:

keep id smesize if smesize<=6

```
. keep id smesize if smesize<=6
invalid syntax
r(198);

.
```

Command



Para lograrlo, se utilizan dos comandos:

keep if smesize<=6

```
. keep if smesize<=6
(0 observations deleted)

.
```

Command

keep id smesize

```
. keep id smesize

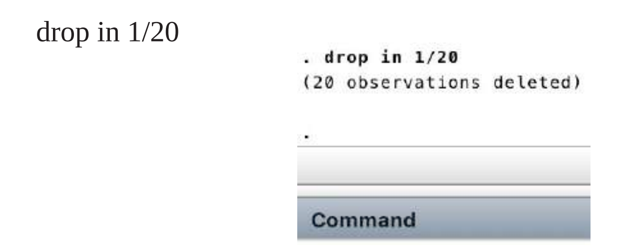
.
```

Command

Variables	
Name	Label
id	ID
smesize	sme size based on personel

También, es posible usar la palabra clave del comando *drop* o *keep*. Por ejemplo, para soltar (*drop*) las primeras 20 observaciones:

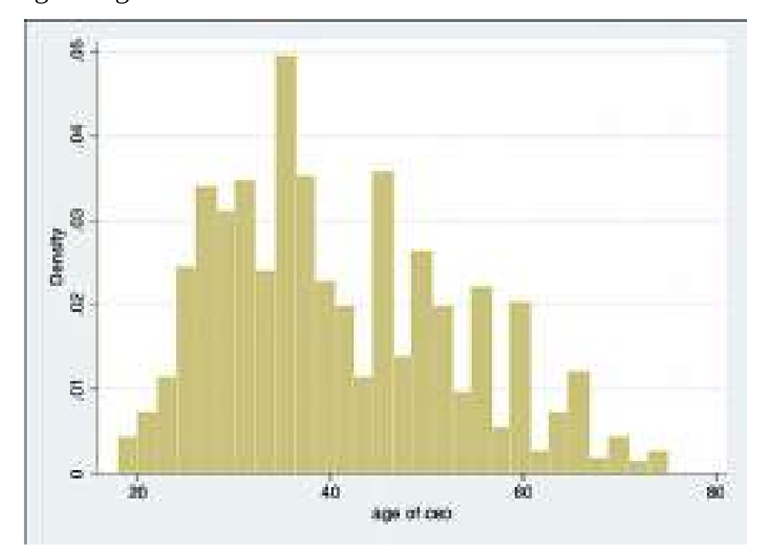
drop in 1/20



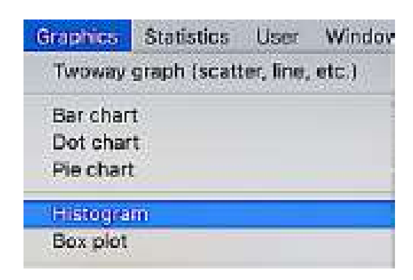
Ejemplo 24. Produciendo gráficos

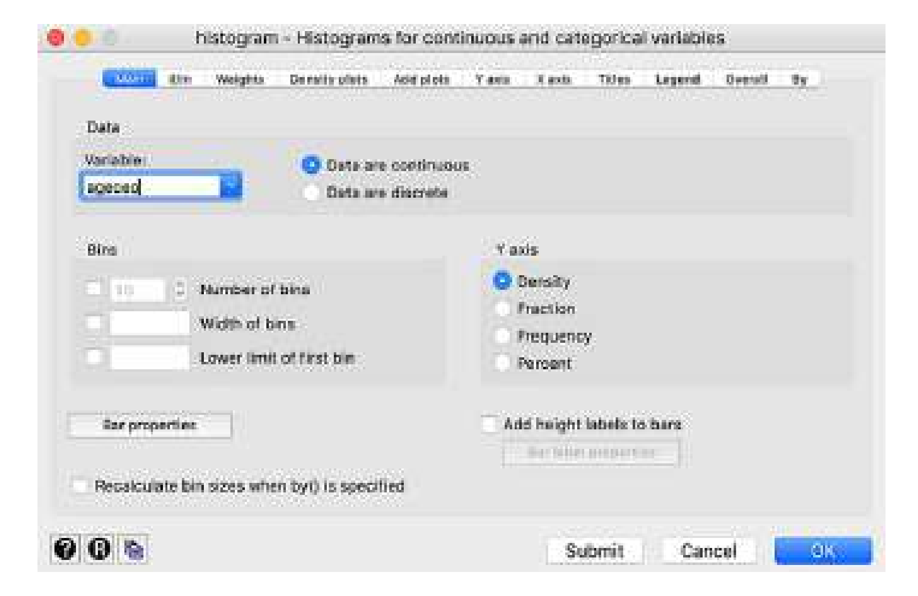
STATA es bastante eficiente en la producción de gráficos básicos, aunque requerirá de considerable experimentación, para producirlos más atractivos. El siguiente comando muestra la distribución de la edad del ceo de las pymes de base tecnológica **(pbt**) en un gráfico de barras (histograma):

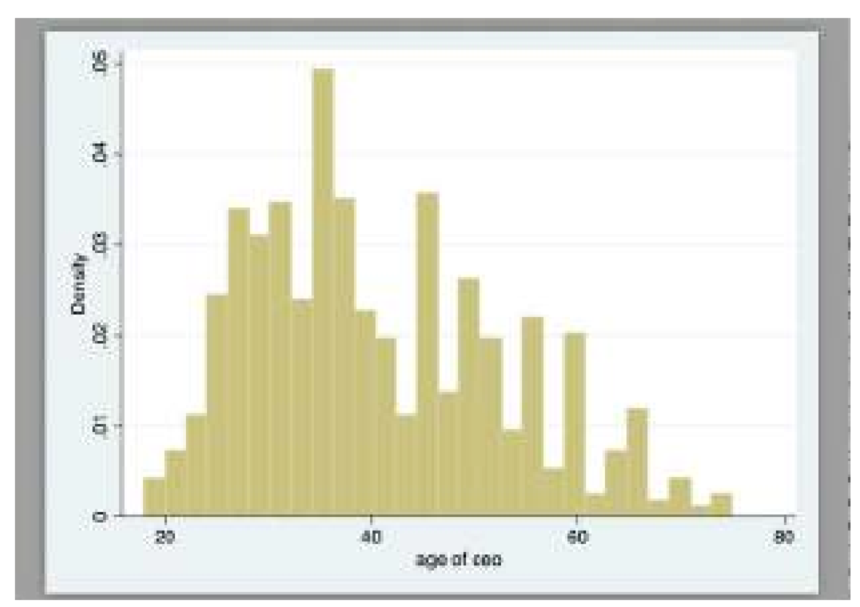
histogram ageceo



En muchos casos, la forma más fácil de producir gráficos es usando los menús; en esto caso, oprima Gráficos y posteriormente, en Histograma y siga las indicaciones.



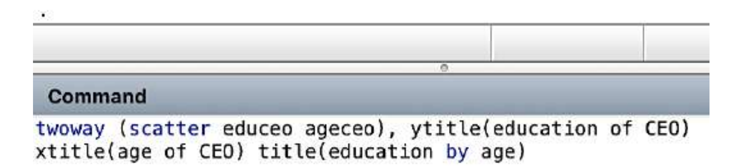

histogram - Histograms for continuous and categorical variables
Main
Bin
Weights
Density plots
Add plots
Y axis
X axis
Titles
Legend
Overall
By
Data
Variable:
ageceo
Data are continuous
Data are discrete
Bins
Number of bins
Width of bins
Lower limit of first bin
Y axis
Density
Fraction
Frequency
Percent
Bar properties
Add height labels to bars
Recalculate bin sizes when by() is specified
Submit
Cancel
OK

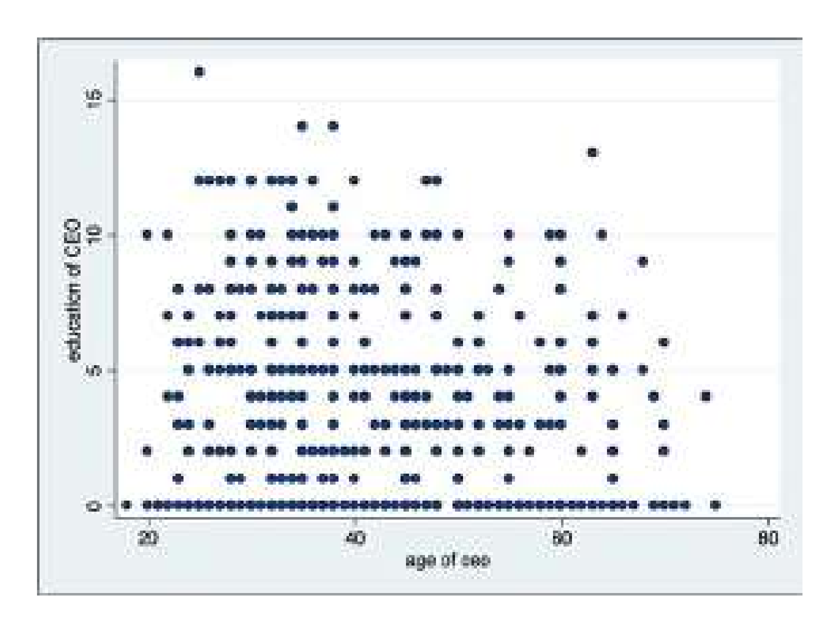

Density
0
.01
.02
.03
.04
.05
20
40
60
80
age of ceo

Una manera fácil de guardar un gráfico es hacer clic derecho sobre él y copiarlo para pegarlo en un Microsoft Word o Documento de Excel.

Aquí hay un comando para un diagrama de dispersión de dos variables:

twoway (scatter educeo ageceo), ytitle(education of CEO) xtitle(age of CEO) title(education by age)

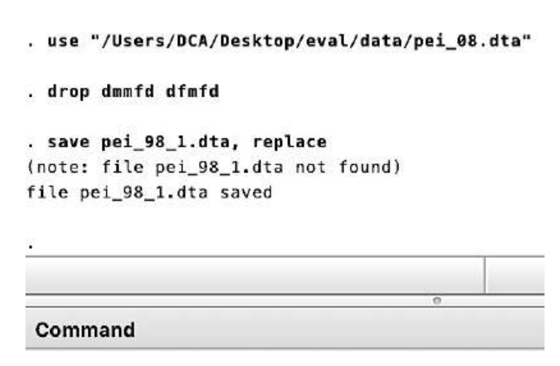




Ejemplo 25. Combinando archivos de datos

La combinación de archivos de datos, se realiza cuando se necesita usar variables que se distribuyen en dos o más archivos. Como ejemplo de combinación, se tiene el archivo el **pei_08.dta**, el cual, se dividirá en dos conjuntos de datos, de tal manera que uno contiene una o más variables que la otra no y entonces y entonces los conjuntos de datos se combinarán para recuperar el **pei_08.dta** original. Así, se requiere abrir el archivo **pei_08.dta**, descartar las variables de participación del programa y guarde el archivo de datos como **pei_98_1.dta**, con los siguientes comandos:

```
use "/Users/DCA/Desktop/eval/data/pei_08.dta", clear
drop dmmfd dfmfd
save pei_98_1.dta, replace
```



Stata/MP 14.1 — pei_98_1.dta

Variables

Name	Label
id	ID
year	year of observation
region	id region
cluster	id cluster
ageceo	age of ceo: years
genceo	gender of ceo: 1=M, 0=F
educeo	education of ceo: years
smesize	sme size based on personnel
smefac	sme assets based on facilities
incinnov	sme incinnov
expopn	sme opn expenditure
expinnov	sme innovation expenditure
exptot	total expenditure
weight	innovation sampling weight
bbaccess	sme broad band accessible all year: 1=Y, 0=N
pcin	proportion of region with internet broadband
raw1	raw1
raw2	raw2
raw3	raw3
raw4	raw4
raw5	raw5
raw6	raw6

Si desea asignarle un nuevo nombre a este archivo (**pei_98_1.dta**) porque no desea cambie el **pei_08.dta** original de forma permanente. Ahora abra el **pei_08.dta** nuevamente. Esta vez, mantenga solo las variables de participación. Guarde este archivo como **pei_98_2.dta.**

```
use "/Users/DCA/Desktop/eval/data/pei_08.dta", clear
keep id dmmfd dfmfd
save pei_98_2.dta,replace
```

```
. use "/Users/DCA/Desktop/eval/data/pei_08.dta", clear

. keep id dmmfd dfmfd

. save pei_98_2.dta,replace
file pei_98_2.dta saved

.
```

Command



Como se observa, se mantiene la *id* de cada registro pyme de base tecnológica, además del resto de los campos. Esto es necesario, porque la combinación requiere al menos una variable de identificación común entre los dos archivos que se van a fusionar. Aquí *id* es esa variable común, entre los dos archivos. Ahora se tienen dos archivos de datos: uno contiene las variables de participación (**pei_98_2.dta**), y el otro no las tiene (**pei_98_1.dta**). Si necesita usar las variables de ambos archivos, tendrá que fusionar los dos archivos. Sin embargo, antes de fusionarlos, necesita asegurarse de que ambos archivos estén ordenados por la variable de identificación. Esto puede hacerse de la siguiente forma:

```
use "/Users/DCA/Desktop/eval/data/pei_98_1.dta", clear
sort  id
save,replace
use pei_98_2, clear
sort id
save,replace
```

```
. use "/Users/DCA/Desktop/eval/data/pei_98_1.dta", clear

. sort  id

. save,replace
file /Users/DCA/Desktop/eval/data/pei_98_1.dta saved

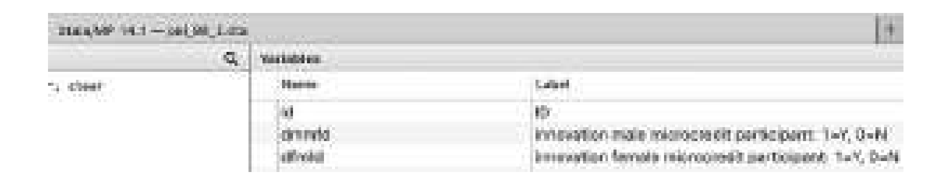
. use pei_98_2, clear

. sort id

. save,replace
file pei_98_2.dta saved

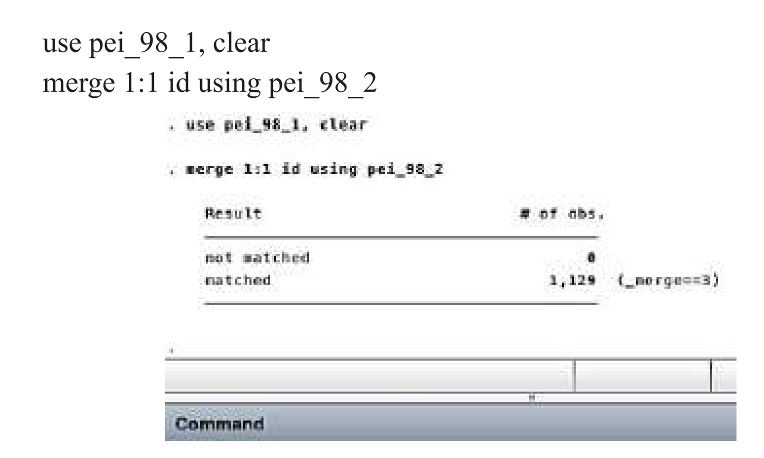
.
```

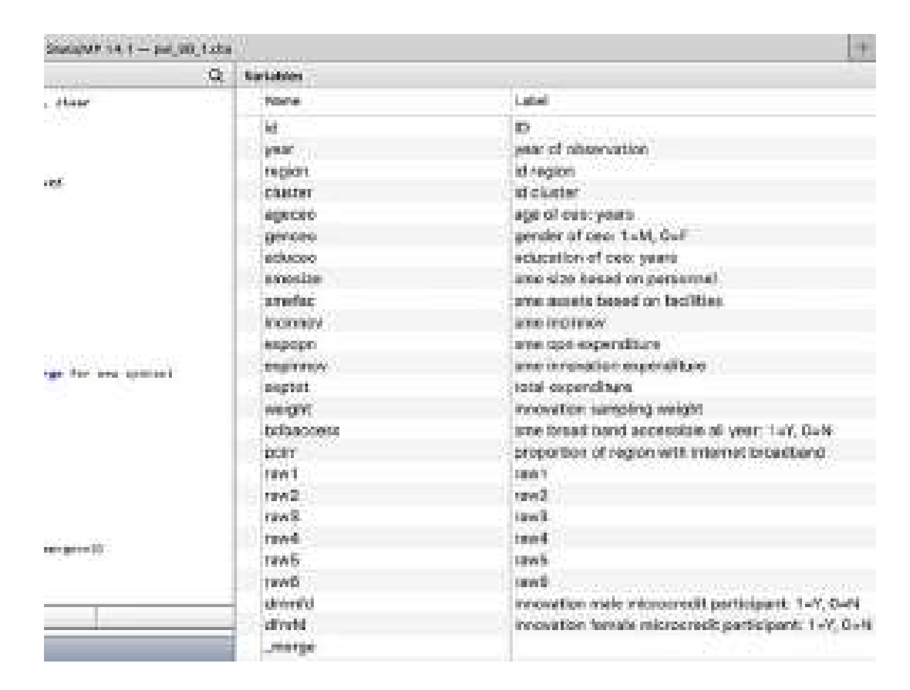
Command



Ahora, se está listo para fusionar los dos archivos. Uno de los archivos, tiene que estar abierto (no importa cual archivo). Abra el archivo **pei_98_1.dta** y posteriormente, combine el archivo **pei_98_2.dta** con él, de la siguiente manera:

use pei_98_1, clear
merge 1:1 id using pei_98_2





_merge
matched (3)
matched (3)
matched (3)
matched (3)
matched (3)
matched (3)

En este contexto, **pei_98_1.dta** se denomina *archivo maestro* (*master file*, es el archivo que permanece en la memoria, antes de la fusión) y **pei_98_2.dta** se llama el archivo de uso (*using file*). Para ver cómo se hizo la operación de fusión, escriba el siguiente comando:

tab _merge

```
. tab _merge

         _merge |      Freq.     Percent        Cum.
----------------+-----------------------------------
    matched (3) |      1,129      100.00      100.00
----------------+-----------------------------------
          Total |      1,129      100.00

.
```

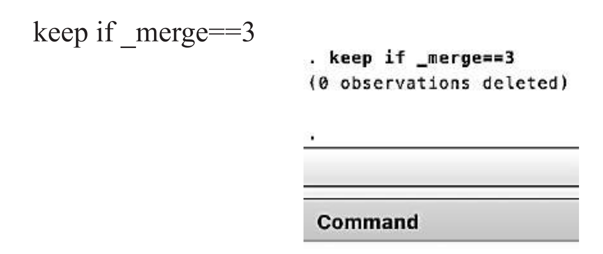
Command

STATA crea una nueva variable *_merge* durante la operación de fusión. Una operación de tabulación a esta variable, muestra los diferentes valores de *_merge* y, por lo tanto, el estado de la operación de fusión. Aunque en este caso, *_merge* solo tiene un valor (3), puede tener hasta tres valores posibles, dependiendo de la naturaleza de la operación de fusión:

- Un valor de 1 muestra el número de observaciones que provienen únicamente del archivo maestro.
- Un valor de 2 muestra el número de observaciones que provienen únicamente del archivo que usa.
- Un valor de 3 muestra el número de observaciones comunes en ambos archivos.

El número total de observaciones en el archivo resultante, es la suma de estos tres frecuencias *_merge*. En este ejemplo, sin embargo, cada observación pyme con base tecnológica, en el archivo **pei_98_1.dta** tiene una coincidencia exacta en el archivo **pei_98_2.dta**, por lo que *_merge = 3* y no 1s o 2s (obviamente, porque los dos archivos se crean a partir de el mismo archivo). Pero en los ejemplos de la vida real, 1s y 2s pueden permanecer después de la fusión. Más a menudo, se requiere trabajar con las observaciones que son comunes en ambos archivos (*_merge = 3*), de la siguiente forma:

keep if _merge==3



Además, es una buena práctica eliminar la variable *_merge* del archivo de datos después de la fusión, con:

drop _merge

Ahora, se tiene un archivo de datos que es idéntico a **pei_98.dta** en el contenido.

Ejemplo 26. Haciendo apéndices en archivos (*Appending data sets*)

Hace agregados a los archivos de datos, es necesario cuando se requiere combinar dos archivos que tienen las mismas (o casi las mismas) variables, pero sus unidades de observación (por ejemplo, las **pbt**) son mutuamente excluyentes Para demostrar la operación *append* (agregar), nuevamente dividirá el archivo **pei_98.dta**. Esta vez, sin embargo, en lugar de descartar variables, retire algunos registros. Así, abra el archivo **pei_98.dta**, borre de 1 a 700 registros y guarde este archivo como **pei_98_1.dta**:

use "/Users/DCA/Desktop/eval/data/pei_98.dta"
drop in 1/700
save pei_98_1.dta, replace
sum

```
. use "/Users/DCA/Desktop/eval/data/pei_98.dta", clear

. drop in 1/700
(700 observations deleted)

. save pei_98_1.dta,replace
file pei_98_1.dta saved

. sum

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
          id |       429    259977.1    35668.11     203163     323103
        year |       429           1           0          1          1
      region |       429     2.11655    .8836158          1          4
     cluster |       429    25.77855    3.562513         20         32
      ageceo |       429    45.82051    12.33387         20         95
-------------+--------------------------------------------------------
      genceo |       429    .9067599    .2911081          0          1
      educeo |       429    2.034965    3.304088          0         14
     smesize |       429    5.407925    2.322992          1         18
      smefac |       429    83.28788    290.4088          1       4208
    incinnov |       429    235740.7     1355599          1   2.42e+07
-------------+--------------------------------------------------------
      expopn |       429    3769.862    1734.835   1149.456   15270.67

Command
```

A continuación, vuelva a abrir **pei_98.dta** pero mantenga las observaciones 1 a 700 y guarde este archivo como **pei_98_2.dta:**

use "/Users/DCA/Desktop/eval/data/pei_98.dta", clear
keep in 1/700
save pei_98_2.dta,replace
sum

```
.
. use "/Users/DCA/Desktop/eval/data/pei_98.dta", clear

. keep in 1/700
(429 observations deleted)

. save pei_98_2.dta,replace
file pei_98_2.dta saved

. sum
```

Variable	Obs	Mean	Std. Dev.	Min	Max
id	700	114809.2	59655.12	11054	203133
year	700	1	0	1	1
region	700	2.217143	.9826785	1	4
cluster	700	11.25286	5.962717	1	20
ageceo	700	46.13	12.89271	18	88
genceo	700	.9085714	.2884237	0	1
educeo	700	2.49	3.568847	0	16
smesize	700	5.234286	2.128752	1	18
smefac	700	72.87588	124.4461	.2	1090
incinnov	700	106447.2	183288.6	1	1733076

Ahora, tiene dos archivos de datos; ambos tienen variables idénticas pero diferentes pymes de base tecnológica (**pbt**). En esta situación, debe agregar (*append*) dos archivos. De nuevo, un archivo tiene que estar en la memoria (cuál, no importa). Abra **pei_98_1.dta** y posteriormente, agregue **pei_98_2.dta.**

. use "/Users/DCA/Desktop/eval/data/pei_98_1.dta", clear
. append using pei_98_2

Tenga en cuenta que no es necesario ordenar los archivos individuales para la operación *append* (agregar), y **STATA** no crea ninguna variable nueva como *_merge* después de la operación de agregar. Se puede verificar que la operación de anexar se haya ejecutado correctamente emitiendo el conteo de **STATA** comando, que muestra el número de observaciones en

el conjunto de datos resultante, que debe ser la suma de las observaciones en los dos archivos individuales (es decir, 1,129).

La siguiente sección, discute el uso de dos tipos de archivos que son extremadamente eficientes en las aplicaciones de **STATA**. Uno almacena los comandos y resultados de para su posterior revisión (*archivos .log*), y el otro, almacena comandos para ejecuciones repetidas más tarde (*archivos do*). Los dos tipos de archivos funcionan de forma interactivamente, lo cual es muy útil para depurar comandos y obtener mayor información de los datos.

Ejemplo 27. Archivos .log (*log files*)

A menudo, se desea guardar los resultados de los comandos **STATA** y también imprimirlos, lo cual es posible realizar del comando *.log*. Dicho archivo, se crea mediante la emisión de un comando *log using* (inicio de sesión) y cerrado por un comando *log close*; todos los comandos emitidos en el transcurso, así como los correspondientes a los resultados (excepto los gráficos), se guardan en el archivo *.log*. Por ejemplo, al utilizar **pei_08.dta** suponga que desea guardar solo la ejecución de la resumen educativo de los jefes por género familiar aquí son los comandos:

```
use “/Users/DCA/Desktop/STATA logfiles/pei_08.dta
log using educeo.log
by genceo, sort: sum educeo
log close
```

```
. use "/Users/DCA/Desktop/STATA log files/pei_00.dta"

. log using educeo.log

      name:  <unnamed>
       log:  /Users/DCA/Library/Mobile Documents/com~apple~CloudDocs/.Trash/BGD_2000_HIES0MP_v01_M_STATA8/educ
> eo.log
  log type:  text
 opened on:  18 Jun 2020, 14:58:24

. by genceo, sort:sum educeo

-> genceo = 0

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      educeo |        104    .8461538    2.314032          0         12

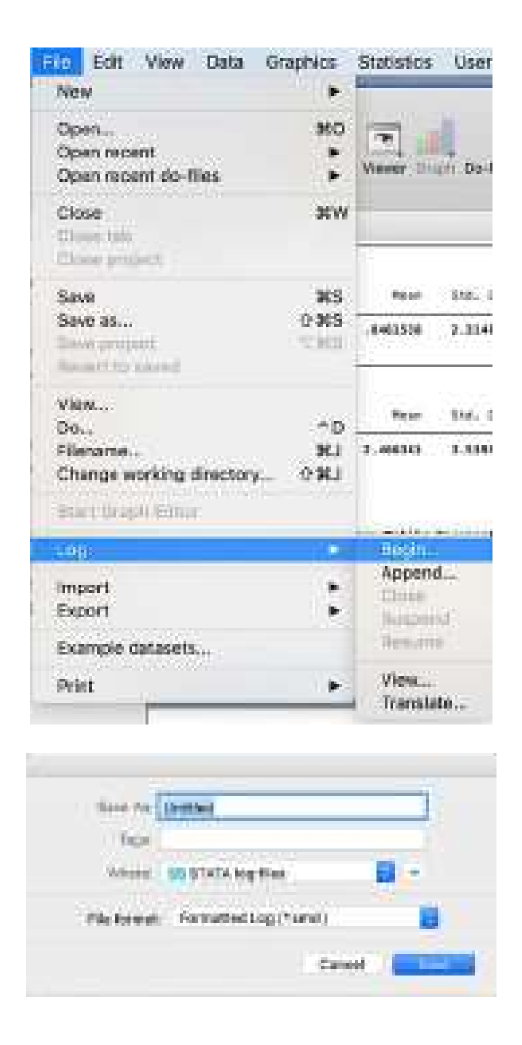
-> genceo = 1

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      educeo |      1,025    2.400341    3.535025          0         16

. log close
      name:  <unnamed>
       log:  /Users/DCA/Library/Mobile Documents/com~apple~CloudDocs/.Trash/BGD_2000_HIES0MP_v01_M_STATA8/educ
> eo.log
  log type:  text
 closed on:  18 Jun 2020, 14:58:24

Command
```

STATA crea un archivo de texto llamado *educeo.log* en el subdirectotrio y guardando los resultados en ese archivo. Si desea que el archivo *.log* se guarde en una carpeta que no sea la carpeta actual, puede especificar la ruta completa, dentrio del archivo *.log*. También puede usar la opción **File→Log→Begin.**



Así, se sugiere tener en cuenta las siguientes opciones:

- Si ya existe un archivo *.log*, puede reemplazarlo utilizando el comando: *log using educeo.log* o también utilizando los comandos: *replace* o *append* a una nueva salida con ingresar *log using educeo.log, append*.

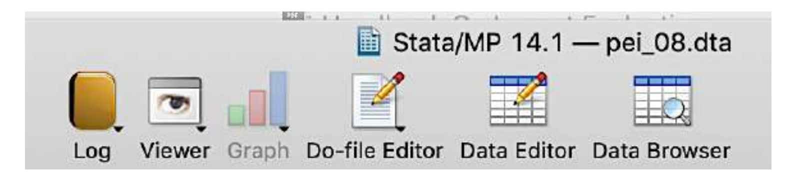
- Si se desea mantener el archivo *.log* existente sin cambios, entonces puede cambiar el nombre de este archivo o el archivo en el comando de creación *.log*.
- Si desea suprimir una parte de un archivo *.log*, puede emitir un comando *lof off* antes de esa parte, seguido de un comando *log on* para la sección que desea guardar. Debe cerrar el archivo *.log* antes de abrir uno nuevo; de lo contrario, recibirá un mensaje de error.

Ejemplo 28. Archivos .do (*.do files*)

Hasta ahora, se ha visto el uso interactivo de los comandos **STATA**, que es útil para la depuración comandos y obtener más información de los datos. Al escribir una línea de comando cada vez, y **STATA** procesa ese comando, muestra el resultado (si lo hay) y espera el siguiente mando. Aunque este enfoque tiene sus propios beneficios, el uso más avanzado de **STATA** implica ejecutar comandos en un *lote*, es decir, los comandos se agrupan y se envían juntos en lugar de uno a la vez. *Si se encuentra usando el mismo conjunto de comandos repetidamente, puede guardar los comandos en un archivo y ejecutarlos juntos siempre que los necesite*. Estos archivos de comando se les conoce como *archivos .do*; son el equivalente **STATA** de las macros. Puede crear archivos .do al menos en tres formas:

- La primera, es escribir simplemente los comandos en un archivo de texto, etiquetado como en nuestro ejemplo: *educeo.do* (el sufijo *.do* es importante), y ejecute el archivo usando *do educeo* en la ventana de comando **STATA** (**STATA** command window).
- La segunda, oprima clic derecho en cualquier lugar de la ventana revisar **STATA** (**STATA** review window) para guardar todos los comandos que fuero utilizados de forma interactiva. El archivo en el que se guardaron, se puede editar y etiquetar y usar como un archivo *.do*.

- Tercero, utilice el *Do-file editor* incorporado en la barra de menú de **STATA**.



Los comandos se pueden escribir en el editor.

Ejecute estos comandos resaltándolos y utilizando el icono apropiado. Con tiempo y práctica, este procedimiento se convierte en una forma muy rápida y conveniente de trabajar con **STATA.**

Aquí hay un ejemplo de un archivo *.do:*

```
log using educeo.log
use pei_98
sort id
save, replace
sort genceo
by genceo: sum educeo
log close
```

Las principales ventajas de usar archivos *.do* en lugar de escribir comandos línea por línea son *replicabilidad y repetibilidad*. Con un archivo *.do*, uno puede replicar resultados que fueron trabajados semanas o meses antes. Además, los archivos *.do* son especialmente útiles cuando los conjuntos de comandos deben repetirse, por ejemplo, con diferentes archivos de datos.

Por ejemplo:

```
*This is a Stata comment that is not executed
/*****This is a do file that shows some very useful
commands used in do files. In addition, it creates a
log file and uses some basic Stata commands ***/
#delimit ;
set more 1;
drop _all;
cap log close;
log using c:\eval\log\try.log, replace;
use c:\eval\data\pei_98.dta ;
describe ;
list in 1/3 ;
list id smesize educeo if genceo==0 & ageceo<45;
summarize smesize;
summarize smesize, detail;
sum smesize educeo [aw=weight], d;
tab genceo;
tab educeo genceo, col row;
tab educeo, summarize(ageceo);
label define genlabel 1 "MALE" 0 "FEMALE";
label values genceo genlabel;
tabulate genceo;
label variable sexhead "Gender of Head: 1=M, 0=F";
save c:\eval\data\temp.dta, replace;
#delimit cr
use c:\eval\data\pei_01.dta
append using temp
tab year
_l_o_g__c_l_o_s_e______________________________________
```

La primera línea del archivo, es un comentario. **STATA** trata cualquier línea que comience con un *asterisco* (*) como comentario y lo ignora. Puede escribir un comentario multilínea utilizando (/ *) como inicio del

comentario, y finalizarlo con (* /). Los comentarios son muy útiles para fines de documentación y deben incluir al menos: *el propósito general del archivo .do y la fecha y hora de la última modificación*. Usted puede incluir comentarios en cualquier parte del archivo *.do*, no solo al principio. Así, de lo contenido, se recomienda tomar en cuenta :

delimit. Por defecto, **STATA** asume que cada comando termina con el comando retorno de carro (es decir, presionando la tecla *Enter*). Si acaso, un comando, es demasiado largo para quedar en una línea, puede extenderlo en más de una línea con este comando. El comando en el ejemplo, dice que un punto y coma (;) finaliza un comando. Cada comando siguiente el comando *delimit* tiene que terminar con un punto y coma. Aunque para este archivo *.do* particular, el comando *#delimit* no es necesario (todos los comandos son lo suficientemente cortos), se hace para explicar el comando con mayor claridad.

set more 1. **STATA**, generalmente muestra los resultados por pantalla a la vez y espera para que el usuario presione cualquier tecla. Pero este proceso, pronto puede convertirse en una molestia si, después de dejar correr un archivo *.do*, tiene que presionar una tecla para cada pantalla hasta que finalice el programa. Este comando, muestra la salida completa saltando página tras página automáticamente.

drop all. Este comando borra la memoria.

cap log close. Este comando cierra cualquier archivo abiert *.log file.*
Si no hay alguno abierto, **STATA** lo ignora.

Ejemplo 29. Archivos .ado (*.ado files*)

Los archivos *.ado* son programas **STATA** destinados a realizar tareas específicas. Muchos comandos **STATA** se implementan como archivos *.ado* (por ejemplo, el comando *summarize*). Para ejecutar dicho programa, simplemente escriba su nombre en la línea de comando. Los usuarios pueden escribir sus propios programas *.ado* para cumplir requisitos especiales. De hecho, los usuarios de **STATA** y los desarrolladores escriben continuamente dichos programas, que a menudo están disponibles a la gran comunidad de usuarios de **STATA** en Internet. Así, se usarán tales comandos a lo largo de los ejemplos sobre diferentes técnicas de evaluación de impacto. Por ejemplo, la técnica de propensión de coincidencia de puntaje (**PSM.***Propensity Score Matching*) se implementa mediante un archivo *.ado* llamado ***pscore.ado.***

findit pscore

STATA responde con una lista de implementaciones de programa *.ado*. Al hacer clic en uno de ellos dará sus detalles y presentará la opción de instalarlo. Cuando Stata instala un *.ado*, también instala los archivos de ayuda asociados.

Glosario

A continuación se ofrece una parte del vocabulario usado en el campo de la evaluación de impacto social en la introducción de innovaciones. Se intentó mantener el enfasis en términos específicos, pero dado que los evaluadores operan en múltiples campos, el lenguaje empleado comprende no solo términos específicos sino también otros que provienen de la evaluación de impacto ambiental y de otros tipos de evaluaciones de impacto, así como también de otros sectores en los cuales esos especialistas se desempeñan. Las descripciones, sin ser definiciones, se incluyen para ayudar a la comprensión de los conceptos. Pese a que muchas descripciones son originales (inspiradas a veces de múltiples fuentes), otras son de uso estándar y de empleo común en varios campos.

Actividad. Medidas adoptadas o trabajo realizado a través del cual los insumos, como los fondos, la asistencia técnica y otro tipo de recursos que se movilizan para generar productos específicos, como el dinero gastado, los libros de texto distribuidos o el número de participantes en un programa de empleo.

Acuerdo de desarrollo. Concepto muy similar al del Acuerdo de Impactos y Beneficios (AIB) pero, en vez de comunitario constituir un acuerdo bilateral entre una empresa y una comunidad, que es lo que el AIB tiende a ser, podría ser iniciado por un gobierno.

Acuerdo de participación. Véase *Acuerdo de impactos y beneficios.*

Acuerdos comunitarios. Véase *Acuerdo de impactos y beneficios* (*Community agreements*).

Acuerdos de impactos y beneficios (IAB). Acuerdos vinculantes entre una empresa y una comunidad (a veces también interviene el gobierno)

que esboza los impactos negativos probables que un proyecto creará, los esfuerzos de mitigación que una empresa llevará a cabo, y el alcance de las contribuciones que la empresa proporcionará a la comunidad en forma de empleos y otros beneficios, como en inversión social.

Adquisiciones locales. Se refiere a las políticas y estrategias deliberadas de las empresas que les permiten aprovisionarse de bienes y servicios a partir de proveedores locales, de modo de intensificar los beneficios del proyecto para la comunidad local.

Afrenta cultural. Acto deliberado que insulta u ofende profundamente, como la violación o profanación de sitios sagrados, o la rotura deliberada de tabúes u otras costumbres culturales significativas.

Agravio (Grievance). Toda inquietud percibida que suscite en un individuo o grupo el sentido de tener derechos o de haber sido agraviado, sobre la base de la ley, de un contrato, de promesas explícitas o implícitas, de prácticas habituales o de conceptos generales de equidad.

AIPP (IAP2). Asociación Internacional para la Participación Pública www.iap2.org.

Alienación. Concepto de las ciencias sociales que refiere a los procesos sociales que alienan a los individuos (los convierte en extranjeros o extraños). El resultado de la alienación es una falta de pertenencia y se experimenta una desconexión, falta de sentido, impotencia y falta de agencia. Así, es un proceso que repercute severamente en el bienestar mental (y a la larga, en la salud física) de los individuos afectados. Se manifiesta en aislamiento social, desesperación, depresión médica y una variedad de otros comportamientos relacionados con la salud.

Análisis de costo-beneficio. Estima los beneficios totales previstos de un programa, en comparación con sus costos totales previstos. Su fin es cuantificar todos los costos y beneficios de un programa en términos monetarios y evaluar si los beneficios superan a los costos. Método económico que evalúa alternativas para un negocio usualmente mediante la determinación del ratio de beneficios con respecto a los costos. En décadas pasadas, se empleaba extensamente en las evaluaciones de impacto, pero no tanto en

las de impacto social porque intenta representar todos los impactos únicamente en términos monetarios.

Análisis de costo-efectividad. Compara el costo relativo de dos o más programas o alternativas de programa en términos de alcanzar un resultado común, como la producción agrícola o las calificaciones de los alumnos en los exámenes.

Análisis de género. Un proceso utilizado para examinar y comprender las implicancias de género de una intervención planeada dirigida a las mujeres como asimismo a los hombres, en el contexto cultural de las comunidades afectadas. Un análisis de género debería observar el sexo y las diferencias de género.

Análisis FODA. Considera las fortalezas, oportunidades, debilidades y amenazas de una organización o comunidad.

Análisis de regresión. Método estadístico para analizar las relaciones entre una variable dependiente (la variable que se debe explicar) y variables explicativas. El análisis de regresión normalmente no es suficiente para capturar los efectos causales. En la evaluación de impacto, el análisis de regresión es una manera de representar la relación entre el valor de un indicador de resultado Y (variable dependiente) y una variable independiente que captura la asignación al grupo de tratamiento o grupo de comparación, mientras se mantienen constantes otras características. Tanto la asignación al grupo de tratamiento y de comparación como las demás características son variables explicativas. El análisis de regresión puede ser univariante (si hay solo una variable explicativa; en el caso de la evaluación de impacto, la única variable explicativa es la asignación al grupo de tratamiento o de comparación) o multivariante (si hay varias variables explicativas).

Análisis de sensibilidad. Trata de la sensibilidad del análisis ante los cambios en los supuestos. En el contexto de los cálculos de potencia, contribuye a comprender cuánto tendrá que aumentar el tamaño requerido de la muestra bajo supuestos más conservadores (como un menor impacto esperado, una mayor variación en el indicador de resultado o un nivel más alto de potencia).

Antiguos terrenos. Hace referencia al proyecto que tiene lugar en un sitio en el que hubo otros proyectos industriales anteriormente y donde es necesario abordar problemas heredados.
Arraigo a un lugar. Hasta qué punto un individuo tiene sentimientos positivos con respecto a su entorno o comunidad local.
Área de influencia. En una evaluación de impacto ambiental (EIA), es el área física (y componentes tales como aire, agua, suelo) sobre la cual un proyecto produce impactos (incluso abiótico, biótico y socioeconómico) causados por un proyecto (y sus actividades asociadas). Por tanto, comprende no solo el área de la superficie de tierra pero también el funcionamiento de todo ecosistema marino y terrestre, cuencas atmosféricas y cuencas hidrográficas o hidrológicas, y todos los agrupamientos sociales de individuos, comunidades, empresas (especialmente pequeñas y medianas empresas), organizaciones y entidades de gobierno. (Véase *Área de influencia social*).
Área de influencia social. Término cuyo significado es prácticamente el mismo que el de *área de influencia* pero que enfatiza los impactos sociales del proyecto. Dada la movilidad de las personas y el alcance de los impactos sociales, el área de influencia social probablemente sea mucho más amplia en área física que el área física de influencia.
Arraigo. Es estar fuertemente integrado en la comunidad local, enraizado, como un árbol que tiene raíces. El arraigo, un componente del apego al lugar junto con el sentido de pertenencia, se refiere a los vínculos sociales de una persona (capital social) con otros miembros de la comunidad, por ejemplo, el tener parientes en ella, tener amigos de larga data, tener amigos íntimos, etc.
Artesanal. Trabajar con las manos. Un artesano es una persona con habilidad para hacer artesanías. A menudo se aplica a mineros de pequeña escala o a pescadores de subsistencia.
Asalariados ingreso fijo. Se refiere a aquellas personas cuyos ingresos (provenientes de trabajos o pensiones) son, es decir que no son objeto de ajuste por cuestiones de inflación o costo de vida. Las ciudades en auge habitualmente tienen una inflación localizada. Aunque las personas vinculadas con el proyecto reciben en pago una cantidad de dinero suficiente

como para hacer frente a los aumentos del costo de vida (que son por lo general la causa de la inflación), mucha gente percibe ingresos que no se ajustan a los niveles locales de inflación. Esta gente incluye no sólo jubilados, sino también personas con salarios que son fijados a nivel nacional, como enfermeros, docentes, policías y otras personas que trabajan en el gobierno.

Asentamientos informales. Asentamientos que se desarrollan informalmente, esto es, sin planificación adecuada y por lo general en violación de los códigos de construcción y planeamiento, usualmente en tierras para las cuales los residentes no tienen títulos de propiedad. A menudo se emplean materiales desechados para construir viviendas temporarias.

Asignación aleatoria o ensayo controlado aleatorio. Método de evaluación de impacto por el cual cada unidad elegible (por ejemplo, un individuo, un hogar, una empresa, una escuela, un hospital o una comunidad) tiene la misma probabilidad de ser seleccionada para ser tratada en un programa. Con un número suficientemente grande de unidades, el proceso de asignación aleatoria garantiza la equivalencia tanto en las características observables como no observables entre el grupo de tratamiento y el grupo de comparación, y así se descarta cualquier *sesgo de selección*. La asignación aleatoria se considera el método más robusto para estimar los contrafactuales y se le suele considerar como la regla de oro de la evaluación de impacto.

Autodeterminación. Uno de los derechos humanos designados que pertenece a grupos (y no a individuos) y que establece que todos los pueblos deben poder determinar libremente su condición política y perseguir libremente su desarrollo económico, social y cultura.

Autoridad competente. Toda persona u organización que tiene la autoridad, capacidad o facultad, asignada por ley o investida, de llevar a cabo una función determinada. En términos de una evaluación de impacto social, para introducción de innovaciones ambientales en general refiere a la autoridad que otorga esta licencia (Véase también *Organismo regulador*).

Beneficiarios. El estado social, económico, psicológico, espiritual o de salud de un individuo o grupo.

Bienestar subjetivo. Experiencia personal de la vida de uno, nivel de satisfacción con la vida de uno, cuán feliz la gente se siente en general sobre su vida en conjunto.
Buena gobernanza. Una interpretación normativa de cómo debería producirse la gobernanza (de cualquier institución), lo cual incluye un compromiso con la responsabilidad, la transparencia, el imperio de la ley, la creación de capacidad, un proceso inclusivo y participativo.
Buena práctica. Lo que se considera habitualmente apropiado y esperado (es decir, convencional más que vanguardista) en un área. En cambio, mejores prácticas quiere decir vanguardista o líder, y por ende adecuado para propugnar, pero no es de esperarse en todas las circunstancias.

Cadena de resultados. Establece la lógica causal del programa explicando cómo se logra el objetivo de desarrollo. Articula la secuencia de insumos, actividades y productos que se espera que mejoren los resultados.
Cálculos de potencia. Método para determinar cuál es el tamaño de la muestra requerida para que una evaluación de impacto estime con precisión el impacto de un programa, es decir: la muestra más pequeña que permitirá detectar el efecto mínimo detectable. Los cálculos de potencia dependen de parámetros como la potencia (o la probabilidad de un error de tipo II), el nivel de significancia, la media, la varianza y la correlación intra-clústers de los resultados de interés.
Calidad de vida. El bienestar general de un individuo, que difiere del estándar de vida en cuanto a que comprende todas las dimensiones no económicas, subjetivas, de la vida.
Capacidad de carga social. Se refiere a la cantidad de personas que un lugar en particular puede sostener. Mientras que la capacidad de carga (ecológica) es un término adecuadamente establecido en ecología, la capacidad de carga social es un término raramente empleado por los profesionales de las ciencias sociales. Los encargados de los parques hablan de capacidad de carga social para referirse a la cantidad de turistas que ellos creen que sus parques pueden recibir, y en el ámbito del turismo existe un concepto de capacidad de carga social como la cantidad de turistas que

una atracción turística específica (sitio, localidad, región, cultura) puede manejar con comodidad.
Capital humano. La educación, destrezas, conocimiento, habilidad para trabajar y salud que, en conjunto, permiten que las personas persigan sus medios de vida objetivos. (Véase también *Capitales*)
Capital social. Los recursos sociales de los cuales se sirven las personas en pos de alcanzar sus objetivos de subsistencia, lo cual incluye sus redes y la amplitud de sus conexiones y sus relaciones de confianza, reciprocidad e intercambio que facilitan la cooperación, reducen los costos de transacción y conforman la base de las redes de seguridad informales. Incluye las instituciones, relaciones, actitudes, valores y valores compartidos y reglas de conducta social que rigen las interacciones entre las personas y contribuyen al desarrollo económico y social.
Capitales. El marco de referencia para considerar la sostenibilidad y el logro de resultados de desarrollo en términos de activos (o capitales), tal como el capital natural, capital humano, capital social, capital financiero, capital manufacturero, y a veces capital político e institucional, y capital cultural y espiritual. Hay varios enfoques que usan los capitales como elemento central, entre los cuales se encuentra el enfoque de los medios de subsistencia sostenibles.
Captura por la élite. Una situación en la cual los recursos que estaban previstos en beneficio de la mayoría de la población son usurpados (capturados) por un pequeño grupo de ricos y poderosos que son parte de la sociedad, una elite económica, política, educativa o étnica.
Carácter irremediable. Daño irreparable, impactos negativos que no pueden ser mitigados ni remediados.
Carácter irreparable. Que no puede ser reparado.
Censo. Empadronamiento total de una población. Los datos censales abarcan todas las unidades de la población. Compárese con muestra.
Cerco perimetral. La delimitación que se extiende alrededor del emplazamiento de un proyecto. Se la utiliza con el propósito de diferenciar entre cuestiones internas y cuestionesexternas. No obstante, aunque esto pueda

resultar obvio en el sentido técnico, no es un término claro, y en términos de cuestiones de derechos sociales y humanos es irrelevante.

CFI. Corporación Financiera Internacional (*International Finance Corporation*), el brazo del Grupo del Banco Mundial para el financiamiento del sector privado. Es de particular importancia ya que sus estándares de desempeño se convirtieron en referencia internacional y son la base de los Principios de Ecuador.

Ciencia abierta. Movimiento cuyo fin es elaborar métodos de investigación más
transparentes, mediante el registro de los ensayos, la utilización de planes de preanálisis, documentación de datos y registros.

Ciudad de crecimiento rápido. Una comunidad, pueblo o ciudad que experimenta un crecimiento excesivamente rápido.

Ciudad industrial. Un asentamiento en el cual la gran mayoría de las personas trabaja para la misma empresa o por lo menos en una de las muchas empresas que operan en el mismo proyecto.

CLPI (FPIC). Consentimiento libre, previo e informado (véase el desarrollo del tema al respecto en el presente documento).

Clúster. También llamado conglomerado. Grupo de unidades que pueden compartir características similares. Por ejemplo, los niños que asisten a la misma escuela pertenecen a un mismo clúster porque comparten el mismo centro escolar, los mismos profesores y viven en el mismo barrio.

Cohesión comunitaria. Se refiere al sentido de armonía en un lugar (área rural, pueblo o ciudad), que se puede establecer por los niveles de: aceptación y valoración de la diversidad social, sentido compartido de pertenencia en todos los grupos, una visión e imagen ampliamente aceptada del lugar, oportunidades en la vida y acceso a servicios razonablemente similares, y relaciones sociales positivas entre las personas de diferentes contextos.

Coincidencia (Matching). Método no experimental de evaluación de impacto que utiliza grandes bases de datos y técnicas estadísticas para construir el mejor grupo de comparación posible para un determinado grupo de tratamiento sobre la base de características observables.

Communitas. Término en latín de las ciencias sociales que significa un fuerte sentido de comunidad, especialmente el que se construye a través de la participación en ritos o celebraciones comunitarias.
Comoditización. Procesos por los cuales la cultura local y los objetos de la cultura local se convierten en mercancías (*commodities*) y por lo tanto tradiciones religiosas, costumbres y festivales del lugar se reducen a conformar las expectativas de aquellos que los adquieren. Es un concepto que proviene de los impactos sociales del turismo, pero puede ocurrir en cualquier situación donde una cultura local entra en contacto con un grupo acaudalado de extranjeros.
Comparación antes y después. También conocida como *comparación previa posterior* o *comparación reflexiva.* Se trata de un seguimiento de los cambios en los resultados para los beneficiarios del programa a lo largo del tiempo, utilizando mediciones antes y después de la implementación del programa o la política, sin utilizar un grupo de comparación.
Comparaciones de inscritos y no inscritos. También conocidas como *comparaciones autoseleccionadas*. Esta estrategia compara los resultados de las unidades que decidieron inscribirse y las unidades que decidieron no inscribirse en un programa.
Complicidad. Palabra empleada en el ámbito de los derechos humanos. Las empresas no deben ser cómplices de violaciones de derechos humanos por parte de terceros. La complicidad comprende cualquiera de las siguientes situaciones: causa o contribuye a la violación de derechos humanos a través del permiso, agravamiento o facilitamiento de aquel; saber o deber haber sabido que las violaciones de derechos humanos probablemente ocurrieran a partir de su conducta; y estar próximo a la violación de derechos humanos ya sea geográficamente o a través de la fuerza, duración o tono de sus relaciones.
Compromiso comunitario. Término que abarca las muy diversas maneras en que las personas interactúan con los procesos de toma de decisiones y participan en estos. Es similar a la "participación pública" y a "intervención pública", por lo que a menudo estos términos se emplean indistintamente. No obstante, el compromiso comunitario es el que se prefiere actualmente

ya que pone énfasis en un grado mayor de participación o compromiso en el proceso de toma de decisiones y mayor respeto por las personas. También se vincula con un discurso diferente y representa el desarrollo de la comprensión en la transición de la consulta al empoderamiento**.**

Comunidad. Concepto de uso común pero disputado que puede aplicarse a diferentes niveles, si bien en general refiere a la agrupación de personas en un lugar determinado y que se presume comparten un sentido de identidad y algunas interacciones de la vida diaria, y tienen algunas instituciones sociales y políticas comunes. Si bien los individuos experimentan impactos sociales a un nivel personal, para la evaluación de impacto social para la introducciónde innovaciones las personas viven, trabajan y juegan en agrupamientos sociales llamados comunidades, lo que por lo tanto los hace el foco principal de las dichas evaluaciones.

Comunidades a ambos lados del cerco perimetral. Las comunidades que se encuentran en la vecindad inmediata del proyecto y que son particularmente vulnerables a los impactos directos del proyecto, tal como ruido, polvo y vibraciones.

Comunidades afectadas. Aquellas comunidades (y comunidades anfitrionas) que reciben el impacto de un proyecto.

Comunidades anfitrionas. Aquellas comunidades que se encuentran cerca de los emplazamientos de los proyectos, que albergan el proyecto y sus trabajadores, dicho de otro modo, las comunidades que reciben el impacto. En la planificación de reasentamientos, se refiere a las comunidades existentes que absorberán las personas que están siendo reubicadas.

Condicionalidades. En asistencia al desarrollo, se refiere a las condiciones que acompañan a un préstamo, alivio de deuda o ayuda bilateral; en licenciamiento ambiental, se refiere a las condiciones de la licencia.

Conflicto latente. Conflicto que está latente, oculto o encubierto en una comunidad. El personal del proyecto podría no notar la presencia de tensión en una comunidad.

Conocimiento indígena. Véase *Conocimiento local*.

Conocimiento local. El conocimiento que la gente de una localidad o comunidad dada ha desarrollado con el tiempo y que seguirá desarrollando.

Hace referencia a la recopilación de hechos y sistemas de conceptos, creencias y percepciones que las personas tienen con respecto al mundo que los rodea. Incluye también la manera en que las personas observan y miden su entorno, cómo resuelven problemas y convalidan la información. Se emplean una amplia variedad de términos: conocimiento indígena, conocimiento técnico indígena, conocimiento tradicional, conocimiento ecológico tradicional o conocimiento medioambiental tradicional y conocimiento de los aborígenes.

Conocimiento tradicional. Véase *Conocimiento local*.

Consentimiento. Otra palabra para expresar que se está de acuerdo con algo. También puede significar tener la facultad de poder dar o retirar la aprobación para un proyecto. Así, un organismo regulador debe determinar las condiciones del consentimiento. En algunas circunstancias, las personas del lugar también podrían tener la habilidad de dar o retirar su consentimiento para un proyecto (véase **CLPI**). El concepto de consentimiento está altamente asociado con la confianza. Un evaluador de impacto social a menudo puede crear consenso para un proceso de evaluación de impacto social para introducción de innovaciones (o para el proyecto), para lo cual muestra que existe un problema que debe abordarse, que una gran variedad de grupos están abordándolo y que el proceso para resolver el problema es justo.

Consentimiento informado. El principio ético más básico para llevar a cabo investigación con humanos; en esencia, que el participante en la investigación tiene derecho a decidir si quiere participar o no, y que para tomar esa decisión debe basarse en información completa sobre la investigación y los posibles riesgos que conlleva. El consentimiento informado a menudo se lo registra mediante la firma del participante en un formulario a tal efecto. Los especialistas en evaluación de impacto social para introducción de innovaciones deben recibir el consentimiento informado de los participantes en la recolección de datos para la evaluación de impacto social para introducción. Uno de los fundamentos de la protección de los derechos de los sujetos humanos. En el caso de las evaluaciones de impacto, requiere que

los encuestados comprendan claramente los fines, procedimientos, riesgos y beneficios de la recopilación de datos en la que se les pide participar.

Consulta pública, intervención pública, participación pública y compromiso comunitario. Si bien a menudo se emplean indistintamente, existen distinciones importantes. Se puede decir que *intervención pública* es un concepto genérico relacionado a los procesos o que implica al público en los procesos de toma de decisiones. *Consulta* implica pedir las opiniones de la comunidad, mientras que *participación* implica realmente que el público colabora en el proceso de toma de decisiones.

Contenido local. Se refiere a la exigencia, expectativa o compromiso de una empresa para asegurar que el valor quede retenido en el ámbito local a través del empleo o del aprovisionamiento.

Contrafactual. Valor que habría tenido el resultado (Y) para los participantes del programa si no hubieran participado en el programa (P). Por definición, el contrafactual no se puede observar. Por lo tanto, debe estimarse utilizando un grupo de comparación. En psicología, refiere a una representación mental o imagen de una trayectoria alternativa, pasada o futura, como manera de concebir otras posibilidades para lo realmente ocurrido. Así, se permite a los individuos procesar sus sentimientos sobre eventos pasados (como en relación a la culpa, remordimiento, arrepentimiento, y preocupaciones de tipo *¿por qué a mí?*, etc.) y también como forma de aprender a partir de las experiencias. Este aprendizaje se puede formalizar en un análisis de escenarios. En el ámbito de una evaluación, sin embargo, tiene un significado diferente: se refiere a la comparación entre lo realmente ocurrido y lo que podría haber ocurrido en ausencia de intervención.

Correlación. Medida estadística que indica hasta qué punto dos o más variables fluctúan juntas.

Correlación intra-clústers. También conocida como “correlación intraclase”. Se trata del nivel de similitud en los resultados o características entre las unidades de los grupos previamente existentes o clusters en relación con unidades de otros clústers. Por ejemplo, los niños que asisten a la misma escuela normalmente serían más similares o estarían más correlacionados

en términos de sus zonas de residencia o antecedentes socioeconómicos, en comparación con niños que no asisten a esa escuela.

Cosmología. La creencia sobre el origen, la historia, la evolución y las leyes culturales relacionados con el cosmos o universo en una cultura particular o sistema mitológico.

Cosmovisión. La perspectiva a través de la cual una persona (y a veces una sociedad o cultura) concibe su mundo; en otras palabras, un marco de referencia cognitivo. Es un término frecuentemente empleado por especialistas en ciencias sociales y especialistas en evaluación de impacto social para la introducción de innovaciones.

Costo de remplazo. Un concepto de la economía y los seguros que se refiere al costo total de reemplazar un activo. La valorización de los activos destruidos por un proyecto con fines de indemnización puede resultar controvertida. Los peritos tasadores de las compañías de seguros hacen uso frecuentemente del valor depreciado de un activo. En el reasentamiento y el desplazamiento inducido por un proyecto, debería cubrirse el costo total de reemplazo de modo de asegurar que no se empeora la situación de las personas.

Cribado. Un proceso de selección para determinar si es necesaria una evaluación de impacto y, si así fuera, qué tipo de evaluación debería realizarse. En un entorno regulatorio, esto se especifica en las reglamentaciones aplicables. Aunque no se emplea siempre de esta manera, el término podría aplicarse de manera más general a la determinación de los requisitos que deben cumplirse, o a los procedimientos que deben seguirse, que están implícitos en los procedimientos de las empresas, la legislación nacional e internacional, y/o las exigencias de los socios financieros, especialmente cuando tienen participación el Banco Mundial, la CFI y los Principios de Ecuador.

Cuenta de depósito en garantía. Se refiere a sumas de dinero retenidas en concepto de garantía por parte de un tercero, de conformidad con los términos de un acuerdo, y que son liberadas una vez que se hayan cumplido las condiciones del acuerdo y/o con el consentimiento de las partes contratantes y/o mediante una orden judicial u otra acción legal.

Cuestiones heredadas. El significado general de herencia es algo que se deja, como cuando una persona muere (lo que deja a otras personas en su testamento), o cuando se jubila de su trabajo (cuando han creado valor o han hecho un desastre); pero también puede referirse a lo que una empresa o proyecto deja. Aunque la palabra pueda ser neutra/de doble interpretación y pueda significar una situación positiva y la contribución a los resultados del desarrollo, en la evaluación de impacto social (EIS), por lo general, hace referencia al desorden (problema) dejado por proyectos anteriores. En el sector extractivo, se refiere mayormente a la polución y la contaminación que queda, tal como el drenaje ácido de las minas. Es importante observar que las cuestiones heredadas de otros proyectos afectan la confianza de la comunidad para aceptar un nuevo proyecto.
Cultura. Aspectos materiales y no materiales de una forma de vida que se moldean y transmiten entre los miembros de una comunidad o de una sociedad mayor. A veces se la denomina creencias, valores, normas, comportamientos, lengua y objetos materiales compartidos que pasan de una generación a la siguiente.
Cumplimiento. Fenómeno que se produce cuando las unidades adhieren a su asignación como parte del grupo de tratamiento o del grupo de comparación. Significa cumplir con la ley y todo reglamento que regule la actividad. En el contexto de una evaluación de impacto, se refiere al grado de cumplimiento de las condicionalidades para el licenciamiento del proyecto. En general, se espera que exista una auditoría periódica o un monitoreo para garantizar el cumplimiento.
Cumplimiento imperfecto. Discrepancia entre el estatus de tratamiento asignado y la condición de tratamiento real. Se produce cuando algunas unidades asignadas al grupo de comparación participan en el programa, o cuando algunas unidades asignadas al grupo de tratamiento no participan.

Datos administrativos. Datos recopilados asiduamente por organismos públicos o privados como parte de la administración de un programa, normalmente con frecuencia periódica y a menudo en el lugar de la prestación de servicios, e incluyen los servicios prestados, los costos y la partici-

pación en el programa. Los datos de monitoreo constituyen un tipo de datos administrativos.

Datos de encuesta. Datos que cubren una muestra de la población de interés. Compárese con censo.

Datos de monitoreo. Datos provenientes del monitoreo del programa que proporcionan información esencial a propósito de la prestación de una intervención, e incluyen quiénes son los beneficiarios y qué beneficios o productos del programa pueden haber recibido. Los datos de monitoreo constituyen un tipo de datos administrativos.

Deber de diligencia. Una obligación de adoptar un cuidado razonable para evitar causar un perjuicio previsible a otra persona o a sus bienes. Un especialista en la evaluación de impacto social para la introducción de innovaciones tiene el deber de diligencia profesional con respecto al cliente y una responsabilidad ética hacia la comunidad, a fin de asegurar que se hayan planteado todas las cuestiones apropiadas.

Debida diligencia. En términos generales, es la investigación que una persona o empresa realizaría antes de celebrar un contrato o antes de realizar una adquisición, especialmente en situaciones en las que pueden existir riesgos. En el ámbito de la evaluación de impacto social para la introducción de innovaciones), refiere prácticamente al mismo concepto excepto que con mayor referencia al Principio 17 de los Principios rectores de las Naciones Unidas sobre las empresas y los derechos humanos que afirma que "*con el fin de identificar, prevenir, mitigar y responder de las consecuencias negativas de sus actividades sobre los derechos humanos, las empresas deben proceder con la debida diligencia en materia de derechos humanos. Este proceso debe incluir una evaluación del impacto real y potencial de las actividades sobre los derechos humanos, la integración de las conclusiones, y la actuación al respecto; el monitoreo de las respuestas y la comunicación de la forma en que se hace frente a las consecuencias negativas*".

Debida diligencia en materia de derechos humanos. Se entiende el objetivo de los Principios Rectores de las Naciones Unidas sobre las Empresas y los Derechos Humanos de que las empresas deben adoptar un proceso de

debida diligencia para garantizar que una acción, transacción o adquisición comercial propuesta no oculte riesgos en materia de derechos humanos (en otras palabras, riesgos para las personas y comunidades, no únicamente riesgos para la empresa). Dado que muchos impactos sociales también son impactos de derechos humanos, los actores afectados son titulares de derechos jurídicos.

Declaración de impacto ambiental (DIA). Documento formal elaborado por una evaluación de impacto ambiental que se presenta ante la autoridad competente.

Declaración de impacto social. El equivalente a una Declaración de Impacto Ambiental en el marco de una evaluación de impacto social para la introducción de innovaciones, un documento formal que se presenta ante una entidad reguladora.

Déficit de gobernanza. Existen muchos tipos de déficits de gobernanza, pero en general se refieren a la diferencia entre la práctica ideal (o al menos lo que debería esperarse normalmente como correcto) y la práctica real de la gobernanza. En la evaluación de impacto y los discursos de empresas y de la sociedad, el déficit tiende a referirse a la falta de observación de las actividades de las empresas multinacionales en los países en desarrollo.

Deliberación. (Y estado deliberativo) Un concepto multidimensional que puede definirse como un diálogo destinado a inducir una reflexión profunda (es decir, un análisis serio) acerca de las opciones y posibilidades, de manera abierta e inclusiva (o sea, sin la intrusión del poder o de la política), y que contemple las inquietudes de todas las partes interesadas.

Denuncia y descrédito. Se refiere a una estrategia que intenta lograr un compromiso o el cumplimiento de las normas o buenas prácticas previstas haciendo públicos los nombres de los infractores o delincuentes. A menudo se la contrasta con *saber y hacer saber*.

Dependencia a un lugar. Similar al apego a un lugar pero se emplea para destacar el alcance de la dependencia o atadura a un lugar determinado, que lo imposibilita de moverse y por lo tanto lo hace vulnerable a cambios.

Dependencia económica. Una situación en la cual una comunidad local o una región dependen fuertemente de una empresa o industria. Dicho de otro

modo, cuando una elevada proporción de la población de la región trabaja para esa empresa o industria.

Derecho consuetudinario. Prácticas y creencias culturales que son parte vital e intrínseca del sistema social y económico de una cultura particular que son tratadas como si fueran leyes y están (semi)legitimadas; esto es, costumbres que se aceptan como requisitos jurídicos o reglas obligatorias de conducta.

Derechos consuetudinarios. Derechos que rigen por costumbre o cultura.

Derechos humanos. Garantías jurídicas universales que protegen a individuos y grupos de personas contra actos y omisiones que afectan las libertades fundamentales, los derechos y la dignidad humana. Las normas de derechos humanos obligan a los gobiernos (principalmente) y a otros órganos de protección a hacer determinadas cosas y les impiden hacer otras.

Desahogo. El proceso de "dejar salir el vapor". En el marco de la EIS, se refiere a situaciones en las que las personas que están enojadas o emotivas pueden expresar sus sentimientos. Las oportunidades de desahogarse son una parte necesaria de un buen proceso de compromiso comunitario.

Desarraigo. Puede referirse a lugares que carecen de un "espíritu de lugar", no son auténticos o están desconectados de su entorno ambiente, o puede referirse a la desconexión que sienten los individuos como consecuencia de su reasentamiento o en razón del grado de cambio ocurrido en su comunidad.

Desarrollo comunitario. Proceso a largo plazo en el cual personas marginadas o que viven en la pobreza aún con esfuerzos para identificar sus necesidades, crear el cambio, ejercer más influencia en las decisiones que afectan sus vidas y trabajan para mejorar la calidad de sus vidas, las comunidades en las que viven y la sociedad que integran.

Desarrollo sostenible. El Informe Brundtland (Nuestro futuro común), define al desarrollo sostenible como el desarrollo que satisface las necesidades del presente sin comprometer la habilidad de las generaciones futuras de satisfacer sus propias necesidades.

Desarticulación social. Proceso mediante el cual se perturban las redes sociales y mecanismos de apoyo dentro de un grupo social. A menudo resulta de la fragmentación que ocurre con un reasentamiento.

Desempeño social. La interfaz entre un proyecto y la sociedad; la configuración de los principios de responsabilidad social de una empresa, sus procesos de sensibilidad social y las políticas, programas y resultados perceptibles en la medida en que se vinculan con las relaciones de la empresa con la sociedad.

Desgaste. El desgaste se produce cuando algunas unidades abandonan la muestra entre una ronda de datos y la siguiente. Por ejemplo, cuando las personas cambian su residencia y no se les puede localizar. El desgaste de la muestra es un caso de falta de respuesta de la unidad. Puede introducir un sesgo en la estimación de impacto.

Desplazamiento. Mientras el reasentamiento es el proceso activo de reubicar personas en razón de la ejecución de un proyecto, el desplazamiento es la experiencia personal y social de pasar por el trastorno de la reubicación, el proceso de pérdida del sentido de pertenencia de cada uno. En los procesos de reasentamiento, el desplazamiento físico refiere a la pérdida de la vivienda como consecuencia de la adquisición de tierras en relación con el proyecto o de restricciones en el uso de las tierras que requieren que las personas afectadas se trasladen a otro lugar. El desplazamiento económico se refiere a situaciones en las que las casas de la gente no se ven afectadas pero pierden otros bienes o el acceso a los bienes (por ejemplo, tierras agrícolas) que generará un perjuicio a los medios de subsistencia y la correspondiente pérdida de ingresos.

Desplazamiento económico. Se refiere al trastorno y los impactos sociales (es decir, desplazamiento) debidos, no ya a la reubicación del lugar de residencia, sino a la pérdida de medios de subsistencia económica, tal como cuando los agricultores pierden sus tierras de cultivo, o cuando la contaminación del agua causa la destrucción de los medios de vida de los pescadores.

Desplazamiento físico. Véase *Desplazamiento*.

Determinación del alcance. El proceso de establecer las principales cuestiones de interés y los actores interesados afectados para determinada intervención planeada.
Determinantes sociales de la salud. Las condiciones económicas y sociales, y su distribución en una población que influyen sobre las diferencias entre individuos y grupos con respecto al estado de salud. Se refiere a los factores de riesgo social en la vida y las condiciones laborales de una persona, más que a los factores individuales (tal como los factores de riesgo del comportamiento o la genética) que influyen en el riesgo de enfermedad o de vulnerabilidad ante una enfermedad o una lesión. Las distribuciones de los determinantes sociales están delineadas por la salud pública y otras políticas, una gobernanza deficiente y ordenamientos económicos injustos donde los que ya son pudientes y sanos pasan a estar en mejores condiciones mientras los pobres que ya tienen más probabilidad de enfermarse se vuelven aún más pobres.
Diferencias en diferencias (DD). También conocido como *doble diferencia o diferencia en diferencia*. La **DD** comparan los cambios en los resultados a lo largo del tiempo entre el grupo de tratamiento y el grupo de comparación. Esto elimina cualquier diferencia entre estos grupos que sea constante a lo largo del tiempo.
Discurso. Implica todo lo que pueda tener que ver con el idioma y la conversación, inclusive todo lo que el uso del idioma conlleva, tal como la construcción activa de pensamientos, identidades y acciones. Es una construcción social que ofrece un conjunto de posibles expresiones acerca de un área determinada y organiza y brinda un marco a la manera como va a hablarse de un tema, un objeto o un proceso en particular. El discurso proporciona descripciones, reglas, autorizaciones y prohibiciones de la acción social e individual.
Diseño cruzado. También denominado diseño transversal. Se produce cuando hay una asignación aleatoria con dos o más intervenciones, lo que permite estimar el impacto de las intervenciones individuales y combinadas.
Diseño de regresión discontinua (DRD). Método de evaluación de impacto cuasi experimental que se puede utilizar en programas que dependen de

un índice límite que determina si los participantes potenciales son elegibles para recibir el programa o no. El umbral de elegibilidad del programa proporciona un punto divisorio entre el grupo de tratamiento y el grupo de comparación. Los resultados para los participantes en un lado del umbral se comparan con los resultados de los no participantes al otro lado del umbral. Cuando todas las unidades cumplen con la asignación que corresponde sobre la base de su índice de elegibilidad, se dice que el DRD es *nítido*. Si hay incumplimiento en el otro lado del umbral, se dice que el DRD es *difuso* o *borroso*.

Diversidad social. Se refiere a la mezcla localizada de grupos sociales e individuos basada en características tales como género, etnia, edad, cultura y contexto económico.

Divulgación. Con frecuencia denominado *divulgación total y franca*, una política de *divulgación abierta* o un *deber de divulgación*, se trata de un término con connotaciones jurídicas y cuasijurídicas que se refiere a la obligación de las partes de una negociación de revelar todo lo que sea considerado pertinente para el tema en estudio (es decir, importancia relativa).

Dominio eminente. Se refiere al poder del Estado para adquirir propiedad privada de manera forzosa. Esto podría hacerse para recuperar (expropiar) tierra para autopistas, aeropuertos, etc. Ocasionalmente, el Estado extiende su poder para habilitar la ejecución de proyectos del sector privado cuando se considera que son de interés nacional.

Economía informal. Se refiere a esa parte de la economía que evita la regulación, imposición o control del gobierno, ya sea porque funciona de manera ilegal o porque funciona en pequeña escala, a base de dinero en efectivo.

Efecto causal. Véase *Impacto*.

Efecto de derrame (*Spillover*). También denominado efecto de contagio. Ocurre cuando el grupo de tratamiento influye directa o indirectamente en los resultados del grupo de comparación (o a la inversa).

Efecto mínimo detectable. El efecto mínimo detectable es un insumo en los cálculos de potencia, es decir, proporciona el tamaño del efecto que una

evaluación de impacto está diseñada para estimar con un determinado nivel de significancia y potencia. Las muestras de la evaluación tienen que ser lo bastante grandes para distinguir al menos el efecto mínimo detectable. Este efecto se determina teniendo en cuenta el cambio en los resultados que justificaría la inversión que se ha hecho en una intervención.

Efecto Hawthorne. Se produce cuando, por el simple hecho de ser observadas, las unidades se comportan de manera diferente.

Efecto John Henry. Acontece cuando las unidades de la comparación se esfuerzan más para compensar que no se les haya ofrecido el tratamiento. Cuando se comparan las unidades tratadas con las unidades de la comparación que se esfuerzan más, la estimación del impacto del programa tiene un sesgo; es decir, se estima un impacto menor del programa en comparación con el impacto real que se obtendría si las unidades de la comparación no realizaran un esfuerzo adicional.

Efecto local promedio del tratamiento (**LATE.** *Local Average Treatment Effect*). Impacto de un programa estimado para un subconjunto específico de la población, como las unidades que cumplen con su asignación al grupo de tratamiento o de comparación en presencia de un cumplimiento imperfecto, o en torno al umbral de elegibilidad cuando se aplica un diseño de regresión discontinua. Por lo tanto, el **LATE** proporciona solo una estimación local del impacto del programa y no debería generalizarse al conjunto de la población.

Efecto de demostración. Las consecuencias que trae a ciertos individuos el observar a otras personas. En la esto puede incluir situaciones en las que los miembros de una comunidad receptora tratan de imitar los estilos de vida, comportamientos, actitudes o idioma de los recién llegados tal como trabajadores o turistas extranjeros. Esto puede generar muchos impactos sociales negativos, entre los que se encuentran el aumento del costo de vida, la frustración y un problemático cambio cultural. También podría eventualmente causar impactos sociales positivos en forma de transferencia de conocimientos entre los contratistas internacionales y las pequeñas y medianas empresas.

Efectos de equilibrio de contexto. Efectos de derrame que se producen cuando una intervención influye en las normas de conducta o sociales en un contexto determinado, como una localidad tratada.
Efectos de equilibrio general. Estos efectos de derrame se producen cuando las intervenciones afectan la oferta y demanda de bienes y servicios y, por lo tanto, cambian el precio de mercado de esos servicios.
Efecto multiplicador. Véase *Multiplicador regional*.
Efecto promedio del tratamiento (**ATE**, por sus siglas en inglés). Impacto del programa bajo el supuesto de cumplimiento total; es decir, todas las unidades que hayan sido asignadas a un programa se inscriben realmente en él, y ninguna de las unidades de comparación recibe el programa.
Efecto tarro de miel. Foco de atracción para migrantes laborales o migración inducida por un proyecto; las personas podrían desplazarse hacia el emplazamiento del proyecto en una tentativa de ser consideradas personas afectadas y por consiguiente con derecho a indemnización, o en busca de trabajo u oportunidades económicas que surjan del proyecto.
Elaboración de perfiles. Proceso de reunir información básica sobre las características de una comunidad y su contexto local en el estado previo al proyecto.
Elefante blanco. Expresión de uso corriente que consiste en una instalación (edificio, etc.) de alto costo y utilidad limitada. Históricamente, la expresión tenía un significado más preciso relacionado con un enorme costo de mantenimiento (un activo que es un pasivo) pero también a la incapacidad de disponer de aquel. Salvo que se tomen las precauciones necesarias en la selección del gasto para la inversión social, muchos proyectos pueden ser elefantes blancos.
Empobrecimiento. El proceso de empobrecerse (que lleva a la pobreza). La pérdida de medios de subsistencia a partir del desplazamiento causado por los proyectos puede conducir al empobrecimiento.
Empoderamiento. Aumentar los recursos y capacidades de diversos individuos y grupos con el fin de comprometer e influir en las instituciones económicas y sociales, e incrementar la responsabilidad de las instituciones públicas. Un proceso participativo que transfiere o coloca la responsabi-

lidad de la toma de decisiones y los recursos en manos de aquellos que se beneficiarán con ellos. Esto puede incluir (i) la creación de capacidad para las organizaciones interesadas; (ii) el fortalecimiento de la capacidad jurídica de las organizaciones interesadas, (iii) la autoridad de las partes interesadas para administrar fondos, contratar y despedir trabajadores, supervisar el trabajo y adquirir materiales; (iv) la autoridad de las partes interesadas para certificar la realización satisfactoria del proyecto y establecer indicadores de monitoreo y evaluación, y (v) el apoyo a iniciativas nuevas y espontáneas de las partes interesadas.

Encuesta de seguimiento. También conocida como encuesta *posterior a la intervención*. Se trata de una encuesta realizada después de que el programa ha comenzado, una vez que los beneficiarios se han beneficiado de él durante algún tiempo. Una evaluación de impacto puede incluir varias encuestas de seguimiento, a veces denominadas encuestas *intermedias* y *finales*.

Enfoque (o marco) de medios de subsistencia sostenibles. Forma de analizar el efecto de los proyectos con respecto a los medios de subsistencia de las personas y comunidades. Emplea los capitales (activos de medios de subsistencia) como base del análisis.

Enfoque basado en derechos. Véase *Enfoque basado en los derechos humanos.*

Enfoque basado en los derechos humanos (lente de los derechos humanos). Marco conceptual y operativo orientado hacia asegurar la promoción y protección de los derechos humanos. Es un enfoque (para la salud, la cooperación en desarrollo, etc.) que procura: (1) ubicar a los derechos humanos y sus principios como el componente central de las acciones; (2) exigir rendición de cuentas y transparencia de parte de los garantes del respeto de los derechos; (3) fomentar el empoderamiento de los titulares de derechos y fortalecer su capacidad para, entre otras cosas, hacer responsables a los garantes referidos; (4) garantizar que la participación significativa de los titulares de derechos en procesos de desarrollo e intervenciones planeadas se reconozca como un derecho intrínseco, no simplemente como mejor práctica; y (5) garantizar el compromiso no discriminatorio de los

titulares de derechos y la priorización de individuos o grupos especialmente vulnerables.

Equidad del impacto. La noción de que los impactos que se producen en una sociedad o generados por un proyecto deben ser compartidos de manera equitativa, de que al menos debería prestarse atención a la justa distribución de los impactos positivos y negativos. Por ejemplo, las trayectorias de vuelo para un aeropuerto deberían ajustarse de modo que la carga de ruido sea compartida, en lugar de que la misma gente reciba todo el ruido.

Equipo de evaluación. Equipo que lleva a cabo la evaluación. Se trata en esencia de una asociación entre dos grupos: un equipo de responsables de las políticas públicas (equipo de políticas) y un equipo de investigadores (equipo de investigación).

EMARF (en inglés, SMART). Específico, medible, atribuible, realista y focalizado. Los buenos indicadores tienen estas características.

Error de tipo I. También conocido como falso positivo. Este error se comete cuando se rechaza una hipótesis nula que, en realidad, es válida. En el contexto de una evaluación de impacto, se comete un error de tipo I cuando una evaluación llega a la conclusión de que un programa ha tenido un impacto (es decir, se rechaza la hipótesis nula de falta de impacto), aunque en realidad el programa no ha tenido impacto (es decir, la hipótesis nula se mantiene). El nivel de significancia es la probabilidad de cometer un error de tipo I.

Error de tipo II. También conocido como falso negativo. Este error se comete cuando se acepta (no se rechaza) la hipótesis nula, pese a que esta última, de hecho, no es válida. En el contexto de una evaluación de impacto, se comete un error de tipo II cuando se llega a la conclusión de que un programa no ha tenido impacto (es decir, no se rechaza la hipótesis nula de falta de impacto), aunque el programa de hecho tuvo impacto (es decir, la hipótesis nula no es válida). La probabilidad de cometer un error de tipo II es 1 menos el nivel de potencia.

ESHIA. Evaluación de Impacto Ambiental, Social y de Salud ("Environmental, Social and Health Impact Assessment").

Espacio deliberativo. Un entorno físico que conduce al estado deliberativo**.**
Espíritu de lugar. Hace referencia a los aspectos únicos, distintivos y preciados de un lugar. Mientras el *sentido de pertenencia* implica los sentimientos personales que un individuo tiene acerca de un lugar, el espíritu de lugar se refiere a las características inherentes al lugar.
Estándar de vida. Indicadores físicos, objetivos del bienestar de un individuo o grupo.
Estimación de tratamiento en los tratados. La estimación del impacto del tratamiento en aquellas unidades que en la práctica se han beneficiado del tratamiento. Compárese con intención de tratar.
Estimador. En Estadística, es una regla utilizada para calcular una característica desconocida de una población a partir de los datos (técnicamente conocido como *parámetro*); una estimación es el resultado de la aplicación real de una regla a una muestra concreta de datos.
Estrategia de salida. En el marco de una EIS o de un proyecto, refiere a la atención que una empresa debe prestar a la manera en que se apartará de las obligaciones a largo plazo que no forman parte de su actividad principal.
Estudio de efectividad. Analiza si un programa funciona en condiciones normales al aumentar la escala. Cuando están adecuadamente diseñados e implementados, los resultados de estos estudios pueden ser más generalizables que en el caso de los estudios de eficacia.
Estudio de eficacia. Analiza si un programa puede funcionar en condiciones ideales. Estos estudios se llevan a cabo en circunstancias muy específicas, a menudo con una fuerte participación técnica de los investigadores durante la implementación del programa. Suelen emprenderse para probar la viabilidad de un programa nuevo. Sus resultados no son generalizables más allá del alcance de la evaluación.
Evaluación. Valoración periódica y objetiva de un proyecto, un programa o una política planificados, en curso o finalizados. Las evaluaciones se utilizan para responder preguntas específicas, a menudo relacionadas con el diseño, la implementación o los resultados.

Evaluación de impacto. Evaluación que intenta establecer un vínculo causal entre un programa o intervención y un conjunto de resultados. Una evaluación de impacto procura responder a la pregunta: ¿cuál es el impacto (o efecto causal) de un programa en un resultado de interés? Proceso de identificar las consecuencias futuras de una acción actual o propuesta.

Evaluación de impacto ambiental (EIA). Proceso formal empleado para predecir las probables consecuencias ambientales (positivas o negativas) de un plan, política, programa o proyecto previo a su implementación, a menudo como parte del procedimiento normativo (licenciamiento ambiental).

Evaluación de impacto cultural. Forma de evaluación de impacto que considera los impactos de un proyecto específicamente sobre la cultura de un grupo social en particular (como un grupo étnico o indígena específico). Consideraría, entre otras cosas, los valores, sistemas de creencias, leyes no escritas, lengua(s), costumbres, economía, relaciones con el medio ambiente local y especies particulares, organización social y tradiciones de la comunidad afectada. Dado que los impactos culturales deberían formar parte de una evaluación de impacto social para la introducción de innovación, la evaluación de impacto cultural es un subcomponente de ésta, pero está estrechamente vinculada a otros impactos sociales con lo cual esa distinción no tendría sentido, excepto en la medida que indique la perspectiva y el propósito de la evaluación de impacto; esto es, la evaluación de impacto social par ala introducción de innovaciones que apunta a comprender los impactos sociales de un proyecto sobre la cultura de la comunidad local.

Evaluación de impacto de igualdad (EqIA). Evalúa aspectos de equidad y discriminación en todas las políticas y estrategias, especialmente con respecto a los grupos vulnerables.

Evaluación de impacto sobre el patrimonio cultural. Proceso de evaluar los impactos posibles de un proyecto propuesto sobre las manifestaciones materiales del patrimonio cultural de una comunidad que incluye sitios, estructuras y restos de valor o importancia arqueológica, arquitectónica, histórica, religiosa, espiritual, cultural, ecológica o estética. Los impactos

sobre el patrimonio cultural intangible se evaluarían en un estudio de impacto cultural.

Evaluación de impacto social. Comprende los procesos de analizar, monitorear y gestionar las consecuencias sociales deseadas y no deseadas, tanto positivas como negativas, de las intervenciones planeadas (políticas, programas, planes, proyectos) y todo proceso de cambio social invocado por tales intervenciones. Su principal objetivo es lograr un ambiente biofísico y humano más sostenible y equitativo. Evaluación de impacto social (forma abreviada) analizar, monitorear y gestionar las consecuencias sociales del desarrollo.

Evaluación de necesidades. Procedimiento sistemático para determinar problemas del cliente o comunitarios y clasificar su importancia como componente del programa de desarrollo. La evaluación de necesidades es el predecesor de los programas de intervención pública actuales.

Evaluación ex-ante Evaluación anticipada. La mayoría de las evaluaciones de impacto se realizan exante, constituyen un pronóstico sobre los probables impactos de una intervención planeada; en otras palabras, acerca de algo que todavía no ha sucedido.

Evaluación ex-post. Evaluación posterior al proyecto. Las evaluaciones ex-post son en realidad evaluaciones de los impactos de un proyecto o política en particular.

Evaluación de proceso. Evaluación que se centra en cómo se implementa y funciona un programa, y que evalúa si se corresponde con su diseño original y documenta su desarrollo y funcionamiento. Compárese con evaluación de impacto.

Evaluación prospectiva. Evaluaciones diseñadas y aplicadas antes de que se implemente un programa. Las evaluaciones prospectivas están incorporadas en los planes de implementación del programa. Compárese con evaluación retrospectiva.

Evaluación retrospectiva. Evaluación diseñada después de que se ha implementado un programa (ex post). Compárese con evaluación prospectiva.

Exclusión social. Procesos que retrasan el logro de la inclusión social y de la integración social, o que conducen a la marginalización de grupos vulnerables.
Experimento de mecanismo. Evaluación de impacto que prueba un mecanismo causal particular dentro de la teoría del cambio de un programa, en lugar de probar el efecto causal (impacto) del programa como un todo.
Expropiar/expropiación. Se refiere a la capacidad de un organismo del sector público, o de una empresa autorizada por el gobierno, para adquirir tierras y otros recursos, sin el consentimiento de los usuarios existentes o residentes. (Véase también *Dominio eminente*).
Externalidad. Un término económico que se refiere a los costos (y por ende a los impactos) que se consideran externos o ajenos a los análisis que se están haciendo, por consiguiente se trata por lo general de los impactos ambiental y social. Un principio de la sostenibilidad es internalizar las externalidades

Facilitación. El proceso de posibilitar que grupos y organizaciones alcancen sus metas, ayudándolos en los procesos que utilizan para colaborar entre sí. Un facilitador es habitualmente independiente, está capacitado y posee experiencia en tareas de facilitación, y cuenta con un repertorio de técnicas para emplear, seleccionadas conforme al objetivo y a los intereses del grupo al que está dirigida esa facilitación. En una evaluación de impacto social para introducción de innovaciones y/o proceso de compromiso comunitario, es un facilitador el que normalmente maneja los procesos inherentes al compromiso comunitario. Las principales aptitudes son la imparcialidad, una capacidad para hacer que las personas se sientan cómodas, una buena comprensión del proceso social y un buen conocimiento de una amplia variedad de técnicas y de cuándo utilizar cada una.
Facipulación. Una palabra inventada que combina las palabras facilitación y manipulación. Se refiere al sentimiento que tienen las personas cuando han intervenido en un proceso de participación comunitaria pero se quedaron con la sensación de haber sido manipuladas.

Factor invariante en el tiempo. Factor que no varía a lo largo del tiempo; es constante.

Factor variante en el tiempo. Factor que varía a lo largo del tiempo.

Falta de rango común. Cuando se utiliza el método de coincidencia (*matching*), la falta de rango común es una falta de superposición entre los puntajes de propensión del grupo de tratamiento, o inscrito, y los del grupo de no inscritos.

Falta de respuesta. Se produce cuando faltan datos o los datos son incompletos para algunas unidades de la muestra. La falta de respuesta de la unidad surge cuando no hay información disponible para algunas unidades de la muestra, es decir, cuando la muestra real es diferente de la muestra planificada. Una forma de falta de respues ta a nivel de la unidad es el desgaste. La falta de respuesta de una entrada se produce cuando los datos son incompletos para algunas unidades de la muestra en un determinado momento del tiempo. La falta de respuesta puede generar sesgos en los resultados de una evaluación si está asociada con la condición de tratamiento.

Falta de respuesta de la unidad. Surge cuando no hay información disponible para un subconjunto de unidades; es decir, cuando la muestra real es diferente de la muestra planificada.

Falta de respuesta de una entrada. Ocurre cuando los datos son incompletos para algunas unidades de la muestra.

Fecha límite. En los procesos de reasentamiento indica la fecha luego de la cual la población no será incluida en la lista de personas identificadas como afectadas por un proyecto, y por consiguiente no tendrán derecho a asistencia e indemnización por reasentamiento.

FIFO. Una sigla referida al uso de trabajadores que normalmente no residen en la vecindad del proyecto y que por lo general son introducidos como trabajadores, habitualmente por determinados períodos de tiempo. Se origina en la expresión en inglés *fly in, fly out* (llevar y traer en avión).

Fondo social. La provisión de financiamiento (quizás, pero no necesariamente en carácter de indemnización) proveniente de un proyecto para ser dirigida a las poblaciones afectadas. Habitualmente los administra la comu-

nidad y se destinan a la promoción de proyectos sociales en beneficio de la comunidad toda. Con frecuencia se calculan como un porcentaje de algo, tal como dólares por megavatio hora de electricidad producida, o millar de toneladas de mineral extraído de una mina. Un fondo social no es un pago en concepto de regalías, aunque los pagos de regalías también podrían ingresar a los fondos sociales.

Fortalecimiento de la capacidad. Proceso coordinado de intervenciones, tales como programas de capacitación usualmente centrados en fortalecer capital humano y en mejorar prácticas institucionales y mecanismos de governanza.

Garante. En el enfoque basado en los derechos humanos, los derechos humanos conllevan tanto derechos (y por ende titulares de los derechos) como obligaciones (y en consecuencia, garantes de esos derechos). En la legislación de derechos humanos, los garantes tienden a ser principalmente Estados, pero también podrían incluir a todos los individuos, en particular las empresas y sus proveedores y contratistas. En el marco del derecho internacional, los Estados asumen obligaciones y deberes de respetar, proteger y cumplir con los derechos humanos. La obligación de respetar significa que los Estados deben abstenerse de restringir o interferir con el ejercicio de los derechos humanos. La obligación de proteger requiere que los Estados protejan a los individuos y grupos contra las violaciones de los derechos humanos. La obligación de cumplir significa que los Estados deben adoptar medidas firmes para facilitar el ejercicio de los derechos humanos básicos. A nivel individual, aunque estemos facultados para ejercer nuestros derechos humanos, también debemos respetar los derechos humanos de los demás.

Garantía ambiental. Un instrumento financiero (a menudo en forma de cuenta de depósito de garantía) que proporciona una garantía que asegure que un proyecto cumplirá con los requisitos de rehabilitación ambiental.

Gastos iniciales. Un término propio de la gestión de proyectos que implica gastar más dinero al comienzo de modo de contar con un mejor diseño y por consiguiente ahorrar dinero en una etapa posterior.

Generabilidad. La medida en que los resultados de una evaluación en un ámbito local serán válidos en otros contextos y en otros grupos de población.
Género. Se refiere a los roles socialmente construidos y asignados a hombres y mujeres. Estos roles se aprenden, cambian con el tiempo y varían ampliamente dentro de una misma cultura y entre culturas diferentes**.**
Gentrificación. El proceso gradual por el cual un lugar (por lo general un barrio céntrico pobre o una pequeña localidad ubicada en un lugar estético) se transforma, dejando de pertenecer a la clase trabajadora y pasando a ser de clase media o clase media alta en cuanto a sus características y composición. Una consecuencia de este proceso es un aumento del valor de las propiedades y del costo del alquiler de las viviendas, lo cual provoca el desplazamiento de los residentes anteriores. Cuando la gentrificación conduce a la turistificación, también puede llevar a un conflicto entre los residentes establecidos (antiguos) y los turistas y/o los nuevos residentes (recién llegados).
Gestión de flujos migratorios. El proceso de manejo de grandes cantidades de personas que llegan a los emplazamientos de los proyectos en búsqueda de oportunidades económicas. (Véase *Efecto tarro de miel)*
Gobernanza. Hace referencia al modo de gobernar de organizaciones, instituciones, empresas y gobiernos. Es el acto de gobernar, el que implica la aplicación de leyes y reglamentos, pero también costumbres, estándares éticos y normas. (Véase también *Buena gobernanza*)
Grupo de comparación. También conocido como grupo de control. Un grupo de comparación válido tendrá las mismas características, en promedio, que el grupo de beneficiarios del programa (grupo de tratamiento), con la única diferencia de que las unidades del grupo de comparación no se benefician del programa que se evalúa. Los grupos de comparación se utilizan para estimar el contrafactual.
Grupo de control. También conocido como grupo de comparación (véase la definición).
Grupo de tratamiento. También conocido como grupo tratado o grupo de intervención. El grupo de tratamiento es el grupo de unidades que es objeto

de una intervención versus el grupo de comparación, que no es objeto de ella.
Grupos marginados. Véanse *Grupos minoritarios* y *Grupos vulnerables.*
Grupos minoritarios. Un término de las ciencias sociales empleado para referirse a agrupaciones sociales que se diferencian o pueden diferenciarse de la mayoría de la sociedad. Por lo general sufren discriminación y subordinación, tienen rasgos físicos y/o culturales que los apartan, y en razón de los cuales son marginados por el grupo dominante, un sentido compartido de identidad colectiva y preocupaciones en común; reglas sociales compartidas acerca de quién pertenece y quién no determina la condición de minoría; y la tendencia a contraer matrimonio dentro del grupo.
Grupos vulnerables. Grupos caracterizados por alguna vulnerabilidad. Si bien esta depende del contexto y puede incluir una amplia gama de grupos, en general el concepto abarca: pueblos indígenas, minorías étnicas, migrantes, personas discapacitadas, personas sin hogar, pobres, quienes luchan contra el abuso de sustancias y personas mayores aisladas.
Guardián. Una persona o institución que controla el acceso a algo. Los guardianes pueden ejercer roles formales o informales. En el contexto de una evaluación de impacto social para la introducción de innovaciones, el concepto indica con frecuencia aquellos individuos que tienen poder para facilitar o impedir el acceso de los consultores a una comunidad específica. Dicho de otro modo, ocupan una posición influyente clave.

Habitabilidad. Los aspectos de un lugar que hace que sus habitantes estén felices de vivir y trabajar allí, y que ofrecen una alta calidad de vida para todos sus habitantes.
Hipótesis. Explicación propuesta de un fenómeno observable. Véase también *Hipótesis nula e Hipótesis alternativa*.
Hipótesis alternativa. Suposición de que la hipótesis nula es falsa. En una evaluación de impacto, la hipótesis alternativa suele ser la hipótesis de que la intervención tiene un impacto en los resultados.
Hipótesis nula. Hipótesis que puede ser falsificada sobre la base de los datos observados. Normalmente, la hipótesis nula propone una posición

general o por defecto. En la evaluación de impacto, la hipótesis nula suele ser que el programa no tiene impacto, es decir: la diferencia entre el resultado del grupo de tratamiento y el grupo de comparación es cero.

Historial de impacto. Se refiere a la experiencia previa que una comunidad ha tenido con otros proyectos. Afecta la manera en que se relacionan con nuevos proyectos y el nivel de confianza que podrían tener. Significa también que podrían existir cuestiones heredadas de las cuales un operador nuevo debe encargarse.

Huella social. Un concepto que intenta ser el equivalente social de la *huella ecológica*, es decir una metáfora con referencia al grado de daño social generado por un proyecto o producto. Es un concepto que no se ve propiciado por los profesionales de las ciencias sociales y por ende no se aplica en la evaluación de impacto social para la introducción de innovaciones, pero algunos profesionales de las ciencias físicas promueven este concepto junto con el de huella ecológica.

IAIA. Asociación Internacional para la Evaluación de Impactos (International Association for Impact Assessment) www.iaia.org.

ICMM. Consejo Internacional de Minería y Metales (International Council on Mining and Metals) www.icmm.com.

Impacto. También conocido como efecto causal. En el contexto de las evaluaciones de impacto, un impacto es un cambio directamente atribuible a un programa, a una modalidad del programa o a innovaciones de diseño. Consecuencia económica, social, ambiental u otra que se puede predecir razonablemente y medir previamente en caso de implementarse una acción propuesta.

Impacto (efecto) indirecto. Ocurre como resultado de otro cambio causado por una intervención planeada. En el ámbito de la evaluación de impacto social para la introducción de innovaciones, un efecto indirecto podrá ser causado por un cambio físico en el medio ambiente. Por ejemplo, la mina podrá aumentar la turbiedad del río, lo que podrá reducir el abastecimiento de peces, lo que puede llegar a disminuir los medios de vida económicos de

los pescadores de ese pueblo. También pueden existir efectos secundarios, efectos de segundo o mayor grado.

Impacto (o efecto) directo. Ocurre como resultado directo de una intervención planeada. También denominado impacto principal o de primer orden. En la evaluación de impacto social para introducción de innovaciones refiere a los cambios sociales e impactos sociales directamente causados por el proyecto en sí, como las molestias que ocasiona a las personas el ruido de maquinaria del proyecto.

Impacto de orden superior. Los impactos sociales indirectos que se producen luego de los impactos de primer orden inmediatamente anteriores en la cadena de impactos que surgen de un proyecto.

Impacto de primer orden. Los impactos que constituyen las consecuencias directas e inmediatas de un proyecto.

Impacto percibido. Algo que se cree que es un posible impacto en vez de algo que se haya establecido como un impacto real. Nótese que los impactos percibidos afectan lo que las personas sienten con respecto al proyecto y cómo se sienten y se comportan en general, por lo que la percepción para ellos es real.

Impacto potencial. Impacto que se ha predicho, y no un impacto real que ya haya ocurrido.

Impacto real. Los impactos sociales que las comunidades sufren efectivamente y no los que se pronostica que ocurrirán.

Impacto residual. Los impactos adversos predichos y que permanecen incluso luego de que se han implementado medidas de mitigación.

Impacto social. Algo que se experimenta o se siente, en el sentido perceptual o corporal a nivel de individuo, unidad social (familia/hogar/colectividad) o comunidad/sociedad. (Véase también *Proceso de cambio social*)

Impactos acumulativos. Los impactos sucesivos, incrementales y combinados de uno o más proyectos (existentes, actuales y futuros proyectos predecibles) sobre la sociedad, la economía o el medio ambiente. Pueden resultar de una agregación o interacción de impactos dentro de un sistema social o ambiental y se definen desde la perspectiva de las personas o del medio ambiente que los experimenta.

Impactos transfronterizos. Se refiere a los impactos, ambientales y sociales, que se trasladan más allá de las fronteras, en general fronteras nacionales, pero se puede decir también cualquier frontera jurisdiccional.
Importancia (evaluación determinación). El acto de asignar alguna forma de prioridad a los problemas que deben tratarse, para su posterior análisis y mitigación. Luego del proceso de determinación del alcance, se procede a evaluar los impactos según su importancia, de conformidad con algunos criterios predeterminados o a través de un proceso deliberativo con una comisión de enlace con la comunidad.
Importancia relativa. Concepto jurídico que hace referencia a si algo es relevante con respecto al asunto por tratar. Por ejemplo, puede referirse a lo que debería divulgarse en una negociación amistosa. También puede referir a lo que podría comprender un informe sobre sostenibilidad.
Inclusión social. Concepto de la justicia social que refiere a un compromiso de políticas y a las estrategias activas del gobierno, en todos sus niveles, y de la sociedad civil para mejorar el acceso de individuos y grupos diversos a oportunidades de desarrollo y a una participación íntegra en la sociedad mediante la eliminación de limitaciones institucionales y la provisión de incentivos y mecanismos.
Indemnización. En los lugares en los que los impactos no se pueden evitar, la indemnización implica restituir a las personas, ya sea individual o colectivamente. Puede ser en forma de pagos en efectivo o de suministro de otras actividades de desarrollo, como la de un hospital, una escuela o una biblioteca pública. Si bien la indemnización puede ser la respuesta que surge de los derechos de propiedad de la comunidad afectada, también la puede realizar el proponente como gesto de buena voluntad o como resultado de una negociación.
Indicador. Variable que mide un fenómeno de interés para el evaluador. El fenómeno puede ser un insumo, un producto, un resultado, una característica o un atributo. Véase también EMARF. Véase también *indicador social*.
Índice de elegibilidad. También conocido como variable forzada. Se trata de una variable que permite clasificar a la población de interés a lo largo de

una línea continua y tiene un umbral o una puntuación límite que determina quién es elegible y quién no lo es.

Indicador social. Medida estadística (variable) empleada para monitorear el cambio en un fenómeno social. La evaluación de impacto social para la introducción de innovaciones identifica indicadores sociales para todos los problemas sociales como temas importantes que necesitan ser monitoreados.

Indicadores clave desempeño. Se establecen para medir el desempeño de una empresa o persona. Los indicadores desempeño de monitoreo por lo general no se consideran indicadores clave de, pero la idoneidad de ese monitoreo debería constituir uno para el directivo responsable de este; y la ausencia de incumplimientos debería ser otro indicador clave de desempeño para una empresa.

Indignación moral. Indignación que provoca la percepción o creencia que se ha violado alguna norma moral, como puede ser la norma de imparcialidad o justicia.

Infraestructura compartida. Se refiere a la infraestructura construida para el proyecto, pero que también se pone al servicio de las necesidades de las comunidades locales. Puede hacer referencia a la generación de electricidad, aprovisionamiento de agua, tratamiento de aguas servidas, como asimismo para puentes, carreteras, líneas ferroviarias, puertos y aeropuertos.

Infraestructura de la comunidad. Servicios e instalaciones públicas y privadas que contribuyen a la calidad de vida general (esto es, salud, transporte, energía, educación, agua, calidad del agua y servicios sanitarios).

Inmigración inducida por un proyecto. Véase *Efecto tarro de miel*.

Innovación social. Un discurso sobre las nuevas maneras de satisfacer las necesidades sociales de las comunidades o de entregar a estas beneficios sociales, a través de un rediseño o la creación de nuevos productos, servicios, estructuras orgánicas, estructuras de gobernanza, políticas, procedimientos y actividades que sean más eficaces que los métodos tradicionales ya existentes del sector público, filantrópicos y dependientes del mercado que se aplican para responder a la exclusión social.

Insumos. Los recursos financieros, humanos y materiales utilizados en la intervención. Intención de tratar (ITT, por sus siglas en inglés). Las estimaciones de ITT miden la diferencia en los resultados entre las unidades asignadas al grupo de tratamiento y las asignadas al grupo de comparación, independientemente de si las unidades de cada grupo recibieron en realidad el tratamiento. Intervención. En el contexto de la evaluación de impacto, se trata del proyecto, del programa o de la política que se evalúa. También conocida como tratamiento.

Integración social. La capacidad de los diferentes grupos de la sociedad para vivir juntos en armonía productiva y cooperativa, y para zanjar diferencias en un marco de interés común en beneficio de todos. La integración social implica justicia para cada individuo y armonía entre los diferentes grupos sociales y países. Significa la integración de grupos desfavorecidos y vulnerables, haciendo que todas las instituciones de la sociedad sean más accesibles para ellos.

Intervención planeada. Proyecto, plan, política o programa. Fundamentalmente, cualquier medida estimada que busca lograr una meta o un resultado definido.

Inversión extranjera directa. Un término comercial referido a las inversiones relacionadas con la obtención de una participación mayoritaria (que por lo general se considera el 10% o más del paquete accionario) en una empresa comercial en un país, por parte de una persona jurídica con sede en otro país.

Inversión social. En el marco de una evaluación de impacto social (EIS), un proyecto debería aportar al desarrollo local facilitando financiamiento para proyectos que contribuyan con los resultados en materia de desarrollo local. La inversión social estratégica se emplea cuando para la empresa que facilita los fondos existe una clara viabilidad comercial.

Inversión social estratégica. Específicamente diseñada para lograr tanto los propósitos estratégicos del proyecto o empresa como las necesidades y aspiraciones de las comunidades locales.

IPIECA. La asociación mundial del sector del petróleo y del gas especializada en cuestiones medioambientales y sociales http://www.ipieca.org/.

Jerarquía de la mitigación. Marco de referencia para planificar medidas de mitigación: evitar, reducir, restablecer e indemnizar. En definitiva, evitar en la fuente, minimizar en la fuente, reducir en el sitio, reducir en el receptor, reparar, indemnizar en especie, indemnizar por otros medios.
Junta de revisión institucional (JRI). Comité nombrado para examinar, aprobar y monitorear la investigación con sujetos humanos. También conocido como Comité de ética independiente o Junta de revisión ética.
Jurado ciudadano. Técnica deliberativa en la que la decisión (quizás sobre seleccionar la mejor alternativa) la toma un panel de (alrededor de 12) ciudadanos legos seleccionados entre el público a quienes se les ha confiado que deliberen sobre temas relevantes en nombre de la comunidad. Si bien la intención es que el panel sea lego y no experto, se espera que aprendan sobre los temas relevantes y consulten a expertos y pidan información según sea necesario. Las decisiones tomadas por un jurado ciudadano tienen mayor legitimidad entre los miembros de la comunidad local que aquellas tomadas en procesos a cargo de expertos.
Justicia social. Se refiere a nociones de justicia y equidad en una sociedad. Se trata de una filosofía acerca del respeto por los derechos humanos, el concepto de que cada uno debería tener la oportunidad de mejorar, y de que deberían tener la posibilidad de participar en decisiones que afectan su propia vida.

Legitimidad. Un concepto que significa que las acciones de una parte son consideradas por un individuo o grupo como convenientes, correctas, apropiadas o al menos aceptables desde el punto de vista normativo de la otra parte. La legitimidad puede interpretarse de varias maneras, entre ellas la legitimidad jurídica, legitimidad política, legitimidad moral y legitimidad social. Por sí sola, al menos en el contexto de una evaluación de impacto social para la introducción de innovaciones, en general por *legitimidad* se entiende *legitimidad social*; la medida en que una acción es aceptable socialmente.
Ley común. La legislación y la interpretación de las leyes que existirían en una comunidad según se manifiestan a través de las decisiones judiciales

(precedentes) de los jueces. Se puede contrastar con el derecho legislado, y está legitimado en países cuyos sistemas jurídicos derivan del sistema inglés. Los sistemas continentales europeos (esto es, derecho romano y francés o código napoleónico) no reconocen la Ley Comun (*Common law*). De alguna manera, es como un derecho consuetudinario.

Ley blanda. Se refiere a los procesos informales de control de las actividades de una empresa, tal como las normas y directivas de la industria. Hace referencia también a acuerdos, convenios, etc. a nivel mundial que influyen en las prácticas de las empresas y a la manera en que son juzgadas o percibidas.

LGBT/LGBTIQ. La comunidad de lesbianas, homosexuales, bisexuales, transexuales, intersexuales y queer. Algunas variantes de esta sigla en inglés incluyen también *en duda* (*questioning*). LGBT, LGBT+ y LGBTIQ son las siglas inclusivas habituales que se emplean en las discusiones sobre la orientación sexual y la identidad de género (SOGI, por sus siglas en inglés). Quienes pertenecen a este colectivo se ven con frecuencia discriminados, y por consiguiente se encuentran entre los grupos marginados y vulnerables de la sociedad.

Licencia ambiental. Permiso administrativo expedido por una autoridad competente mediante el cual se autoriza a un operador de una actividad productiva u obra de infraestructura a llevar a cabo acciones para las cuales se solicita la licencia, si bien quizás dependiendo de si se cumplan ciertas condiciones operativas, se respeten ciertos límites de uso y se implementen ciertas medidas para contener, minimizar y evitar cualquier impacto social o ambiental que la actividad o la obra puedan causar.

Licencia social para operar. Expresión popular que implica que la aceptación por parte de la comunidad también es necesaria para el éxito de un proyecto.

Licencia social para operar y crecer. Variante de *licencia social para operar* que pone énfasis en que la aceptación de todos los actores también es necesaria para que el negocio se amplíe.

Línea de base. Situación previa a una intervención, con respecto a la cual se puede valorar el progreso o se pueden hacer comparaciones. La línea

de base se recopila en forma previa a la implementación de un programa o política para observar la situación antes. La disponibilidad de datos de línea de base es fundamental para documentar el equilibrio en las características anteriores al programa entre los grupos de tratamiento y de comparación. Los datos de línea de base son necesarios para algunos diseños cuasi experimentales. Los datos de una serie de indicadores seleccionados que se miden cerca del inicio del proyecto y que permiten hacer un monitoreo de un cambio en el tiempo. Se convierten en un punto de referencia, junto a otros valores de referencia, frente a los cuales se pueden comparar situaciones futuras. Si bien la línea base original refiere a un punto específico en el tiempo, el perfil comunitario debería destacar tendencias en el zona del proyecto para que se pueda comparar entre lo que probablemente hubiera ocurrido con y sin ese proyecto.
Línea base social. Véase *Línea de base.*
Lugar. Espacio geográfico que tiene un significado (sentimiento negativo tanto como positivo) para individuos o grupos por experiencias personales y/o por las relaciones que estos tienen con aquel.

Mapa mental. Un diagrama que es una manera de presentar ideas y pensamientos; la creación de mapas mentales es un proceso de organización de información e ideas.
Marco de políticas de reasentamiento. Una política o guía operativa para el proyecto acerca de la manera en que deberán abordarse las cuestiones de la adquisición de tierras, reasentamiento, indemnización y recuperación de los medios de vida durante todo el transcurso del proyecto.
Marco muestral. Lista exhaustiva de las unidades de la población de interés. Se requiere un marco muestral adecuado para asegurar que las conclusiones a las que se arribe a partir del análisis de una muestra se puedan generalizar a toda la población. Las diferencias entre el marco muestral y la población de interés crea un sesgo de cobertura. Ante la presencia de dicho sesgo, los resultados de la muestra no tienen validez externa para toda la población de interés.

Marginación. Procesos sociales y económicos que empeoran la situación de los grupos minoritarios o vulnerables.

Mecanismo comunitario de reclamos (Community grievance mechanism). Específicamente diseñado para que accedan a este los miembros de las comunidades afectadas por un proyecto.

Mecanismo de reclamos. Un proceso formal, judicial o extrajudicial para el tratamiento de reclamos, al que pueden acceder individuos, trabajadores, comunidades y entidades de la sociedad civil que han sido perjudicados, o sienten que lo han sido, por las actividades de un proyecto o empresa.

Medio ambiente. Un concepto muy vago que se define de diferentes maneras en diferentes circunstancias. En algunas jurisdicciones, incluye los ecosistemas y las partes que los constituyen, incluidas las personas y las comunidades; recursos naturales y físicos; las calidades y características de las ubicaciones, lugares y áreas; y los aspectos sociales, económicos, culturales, estéticos y patrimoniales de todos estos elementos. En otras jurisdicciones, el medio ambiente se refiere solamente a los elementos biofísicos tal como el agua, el aire, el suelo, la flora y la fauna.

Medio de subsistencia. Un medio de subsistencia comprende las capacidades, bienes (véase *Capitales*) y sostenible actividades requeridas para un medio de subsistencia. Un medio de vida es sostenible cuando puede hacer frente y recuperarse de tensiones y conmociones y mantener o mejorar sus capacidades y recursos tanto ahora como en el futuro, sin socavar la base de los recursos naturales.

Medios de subsistencia. Hace referencia al modo de vida de una persona o de un hogar y a la manera en que se ganan la vida, en particular el modo en que cubren las necesidades básicas, por ejemplo, sus alimentos, agua, abrigo y vestimenta, y viven en comunidad. (Véase *Medios de subsistencia sostenibles*).

Medios de subsistencia alternativos. También llamados oportunidades económicas alternativas, se refieren al proceso de identificar, seleccionar y desarrollar una variedad de actividades generadoras de ingresos para reemplazar o aumentar las actividades corrientes que proporcionan medios de subsistencia a las personas afectadas por un proyecto. Esto tiene parti-

cular importancia en el caso de un desplazamiento económico o físico, pero también pueden integrar los programas de creación de beneficios o de inversión social del proyecto.

Mejoramiento. Medidas tomadas en el diseño y las fases subsiguientes de los proyectos de modo de garantizar la obtención de una amplia variedad de resultados directos e indirectos en materia de desarrollo.

Mejores prácticas (*Best practice*). El conjunto de lineamientos, ética, ideas, procedimientos y métodos que representan la mejor (más adecuada) forma de actuar en un sector determinado. Si bien la mejor práctica sería la indicada por una asociación profesional, en general el concepto es vago y refiere a un punto de referencia que se aspira alcanzar.

Mentalidad asocial. Actitud según la cual las personas no cuentan, o que los problemas sociales no son importantes y no es necesario considerarlos.

Mercado emergente. Otra forma de referirse a un país en desarrollo.

Método cuasi experimental. Métodos de evaluación de impacto que no dependen de la asignación aleatoria del tratamiento. Las diferencias en diferencias, el diseño de regresión discontinua y el de coincidencia (*matching*) son ejemplos de métodos cuasi experimentales.

Método de control sintético. Un método de coincidencia (*matching*) específico que permite estimar el impacto en contextos donde una única unidad (como un país, una empresa o un hospital) es objeto de una intervención o es expuesto a un suceso. En lugar de comparar esta unidad tratada con un grupo de unidades no tratadas, el método utiliza información sobre las características de la unidad tratada y las unidades no tratadas para construir una unidad de comparación sintética o artificial, ponderando cada unidad no tratada de tal manera que la unidad de comparación sintética se parezca todo lo posible a la unidad tratada. Esto requiere una larga serie de observaciones a lo largo del tiempo, tanto de las características de la unidad tratada como de las unidades no tratadas. Esta combinación de unidades de comparación en una unidad sintética proporciona una mejor comparación para la unidad tratada que cualquier unidad no tratada individualmente.

Métodos mixtos. Enfoque analítico que combina datos cuantitativos y cualitativos.

Microfinanzas. Variedad de servicios bancarios (en general, la provisión de pequeños préstamos) para asistir a personas de bajos ingresos (en particular mujeres) para establecer pequeñas empresas. Los contratos de microfinanzas son muy importantes porque proveen fondos a personas que de otra manera no acceden al capital debido al pequeño valor del préstamo requerido, sus bajos ingresos, la falta de garantía y antecedentes de préstamos muy pobres o no existentes.
Miedo al delito. Se refiere a un miedo anormal que la gente siente ante la posibilidad de ser víctima de un delito. Un cambio rápido en una comunidad puede incrementar el miedo de las personas a sufrir un delito, en modo tal que su miedo sea desproporcionadom con respecto a la posibilidad real del delito. Es un estado muy debilitante, porque modifica el comportamiento de las personas, afecta su sentimiento con relación a su comunidad y afecta su bienestar general.
Minería de datos. Práctica de manipular los datos en busca de resultados concretos.
Mitigabilidad. Capaz de ser mitigado.
Mitigación. Proceso de concebir e implementar procesos, procedimientos y/o cambios a una intervención planeada con el fin de evitar, reducir o minimizar, o para compensar, impactos que probablemente ocurran.
Monitoreo. Proceso continuo de recopilar y analizar información para evaluar el desempeño de un proyecto, un programa o una política. El monitoreo suele hacer un seguimiento de los insumos, actividades y productos, aunque ocasionalmente también incluye los resultados. Se utiliza para fundamentar la gestión y las decisiones diarias. También se puede emplear para hacer un seguimiento del desempeño en relación con los resultados previstos, establecer comparaciones entre programas y analizar las tendencias a lo largo del tiempo. Puede referir a un proceso de control de cumplimiento de las condiciones necesarias para obtener el consentimiento para llevar a cabo una intervención planeada, pero en general hace referencia al proceso de prueba continuo que determina que no haya impactos no anticipados.

Monitoreo participativo. La participación de actores en actividades de monitoreo y la verificación de información para garantizar la legitimidad del proceso de monitoreo y del proyecto en su conjunto.

Muestra aleatoria. Muestra extraída a partir de un muestreo probabilístico, por lo cual cada unidad en el marco muestral tiene una probabilidad conocida de ser extraída. Seleccionar una muestra aleatoria es la mejor manera de evitar una muestra no representativa. El muestreo aleatorio no debería confundirse con la asignación aleatoria.

Muestra conglomerada. Una muestra compuesta de clústers.

Muestra estratificada. Se obtiene dividiendo la población de interés (marco muestral) en grupos (por ejemplo, hombres y mujeres) y posteriormente, definiendo una muestra aleatoria en cada grupo. Una muestra estratificada es una muestra probabilística: todas las unidades de cada grupo (o estrato) tienen la misma probabilidad de ser asignadas. Siempre que todos los grupos sean lo bastante grandes, el muestreo estratificado permite elaborar inferencias a propósito de los resultados no solo a nivel de la población sino también dentro de cada grupo.

Muestra. En Estadística, una muestra es un subconjunto de una población de interés. Normalmente, la población es muy grande, lo cual hace impracticable o imposible realizar un censo o un registro completo de todos sus valores. En cambio, los investigadores pueden seleccionar un subconjunto representativo de la población (utilizando un marco muestral) y recopilar estadísticas sobre la muestra. Estas se pueden utilizar para hacer inferencias o para extrapolar a la población. Este proceso se conoce como muestreo. Compárese con censo.

Muestreo. Proceso por el cual las unidades se extraen del marco muestral creado a partir de la población de interés. Se pueden utilizar diversas alternativas de procedimientos de muestreo. Los métodos de muestreo probabilístico son los más rigurosos, ya que asignan una probabilidad bien definida a cada unidad que será extraída. El muestreo aleatorio, el muestreo aleatorio estratificado y el muestreo conglomerado son métodos de muestreo probabilístico. El muestreo no probabilístico (por ejemplo, el muestreo intencional o por conveniencia) puede generar errores de muestreo.

Muestreo probabilístico. Proceso de muestreo que asigna una probabilidad bien definida a cada unidad que será extraída de un marco muestral. Incluye el muestreo aleatorio, el muestreo aleatorio estratificado y el muestreo de clusters.
Multiplicador regional. Término económico que refiere al ratio de la magnitud del efecto total sobre la economía regional de un estímulo inicial específico (como un proyecto) relativo a la magnitud del impacto directo; de otro modo, el alcance de la amplificación de una inversión local en una economía local.
Mundo de la vida. Un concepto (en alemán, *Lebenswelt*) de las ciencias sociales que se refiere a las experiencias vividas por las personas y a sus vidas cotidianas. Implica tener en cuenta su perspectiva en el análisis o narrativa que se presente.

NIMBY. Acrónimo que significa *no en mi patio trasero*" también se lo conoce en español como **SPAN** o *sí, pero aquí no*. Califica a la reacción que se recibe a menudo a la hora de decidir la localización de usos de suelos localmente no deseados **(LULU**, por su acrónimo en inglés), como aeropuertos, vertederos, centrales eólicas, etc.
Norma. Estándar, convención o resolución que se emplea como referencia u objetivo.
Norma de desempeño. Concepto genérico que articula el estándar previsto de una práctica o un logro. En el contexto de una evaluación de impacto social para introducción de innovaciones, se refiere normalmente a las Normas de Desempeño sobre Sostenibilidad Ambiental y Social de la Corporación Financiera Internacional (**CFI**).
Normalización. Intento por evitar que se desarrolle una ciudad industrial o de que se cree un sistema de valores de una ciudad de crecimiento rápido, o por darle a una ciudad industrial una vida comunitaria más normal.
Normativo. En una perspectiva o juicio normativo se afirma como tienen que ser las cosas, o como deberían ser, como las valoramos, qué es bueno o malo y qué medidas están bien o mal; a menudo en referencia a un principio

ético, o se puede decir en referencia a un código o estándar internacional que va más allá de los requisitos jurídicos (si fuera requerido jurídicamente, no sería un requisito normativo).

Oficina del Asesor en Cumplimiento/Defensoria. Oficina independiente (depende directamente del Presidente de la Corporación Financiera Internacional (**CFI**)/Grupo del Banco Mundial) que responde a reclamaciones de aquellos afectados por proyectos financiados por la **CFI**. Únicamente considera si la **CFI** ha seguido los procedimientos apropiados.

ONG de control. Una organización no gubernamental (**ONG**) cuya misión u objetivos en particular incluyen una función de monitoreo para supervisar las actividades de empresas u otras organizaciones. Estas **ONG** de control tienden a emplear la estrategia de denuncia y descrédito para fomentar el uso de mejores prácticas industriales.

Organismo regulador. (Autoridad reguladora, entidad reguladora, autoridad competente) Es la autoridad pública o entidad gubernamental encargada de la toma de decisiones y supervisión, con función fiscalizadora o de control, sobre un área de actividad humana. En el marco de una evaluación de impacto social para la introducción de innovaciones, puede ser la Dirección de Medio Ambiente o la Oficina de Planificación. El organismo es responsable por determinar la aceptabilidad de una Declaración de Impacto Ambiental o de un informe de Evaluación de Impacto Social y por expedir la licencia para llevar adelante el proyecto.

Paradigma. Una serie de prácticas, métodos, teorías y entendimientos respaldados por una disciplina científica (que la define) en un determinado momento de la historia. En otras palabras, la visión del mundo dentro de un área de investigación; esto es, lo que el conjunto de académicos y profesionales que ejercen en un área particular de investigación consideran que es buena práctica corriente.

Participación. Proceso a través del cual los actores influyen y controlan las iniciativas de desarrollo y las decisiones y recursos que las afectan. Proceso

que puede mejorar la calidad, efectividad y sostenibilidad de proyectos y fortalecer la propiedad y compromiso del gobierno y los diferentes grupos interesados.

Patrimonio cultural. El legado de objetos materiales y las cualidades intangibles de un grupo o sociedad que provienen de generaciones pasadas, mantenidas en el presente y concedidas para el beneficio de futuras generaciones. El patrimonio cultural comprende la cultura tangible (edificios, monumentos, libros, artesanías y objetos), cultura intangible (folclor, tradiciones, lengua y saber tradicional) y patrimonio natural (paisajes significativos desde el punto de vista cultural, hábitats importantes para la vida silvestre y biodiversidad).

Patrimonio cultural intangible. Véase *Patrimonio cultural.*

Patrimonio vernáculo. La palabra *vernáculo* remite por lo general al idioma cotidiano de la gente común en un lugar determinado. De manera similar, *patrimonio vernáculo* hace referenciaal patrimonio cultural que atañe a la vida de la gente común o que es específico de pequeños grupos que comparten un idioma o una serie de experiencias comunes.

Perfil comunitario. Descripción de las comunidades que probablemente se vean afectadas por una intervención planeada.

Perfil social. Véase *Perfil comunitario*.

Periodo de recepción de comentarios públicos. En procesos regulatorios de evaluación de impactos en general se exige que la Declaración de Impacto completa esté disponible para permitir que el público comente al respecto durante un periodo de tiempo designado (esto es, 30, 45, 60 o 90 días).

Personas afectadas. Quienes residen en zonas aledañas y escucharán, verán, sentirán u olerán el proyecto propuesto. Se las reubica ya sea voluntaria o involuntariamente. Tienen un interés en el proyecto o en cambios de política (ya sea que residan o no en zonas de influencia primaria o secundaria) y también en los recursos potencialmente afectados. Es probable que usen regularmente la tierra afectada o que las afecte el influjo de residentes zafrales, temporarios o permanentes asociados con el proyecto.

Personas afectadas por el proyecto (PAP). Término de la **CF**I/Banco Mundial que puede a veces significar cualquier persona adversamente afectada por un proyecto, pero a veces se refiere principalmente a personas que deben reasentarse en otro lugar o que de otro modo han sido desplazadas como resultado de un proyecto.

Personas interesadas y afectadas. Grupos, organizaciones o individuos que creen que una acción podría afectarlos o que se encuentran implicados en el proceso de decisión (también denominados partes *interesadas*. Véase también *Personas afectadas*)

Pertenencia. El sentido de pertenecer a un grupo social, que es una necesidad emocional humana importante. Muchas veces, como consecuencia de los proyectos, el sentido de pertenencia se reduce, ya sea por los cambios físicos y sociales que ocurren o la presencia de personas nuevas, pero también por los procesos que ocurren y que inducen a la alienación.

PGIS. Plan de Gestión del Impacto Social.

PIMBY. Acrónimo que en inglés significa "por favor, en mi patio trasero", corresponde a la actitud contraria a **NIMBY**, es decir, obtener una licencia social para operar de forma que la gente desee que haya proyectos en las inmediaciones.

Plan. Estrategia para alcanzar objetivos identificados o un programa de implementación.

Plan de acción de reasentamiento. Elaborado en línea con el marco de políticas de reasentamiento, consiste en una estrategia detallada en cuanto a la manera en que se llevará a cabo un proceso de reasentamiento específico. Detalla los procesos de relevar las condiciones iniciales, consultar a las personas afectadas, y brinda una estrategia detallada para: (i) minimizar o evitar el reasentamiento; (ii) indemnizar por las pérdidas; (iii) reubicar y reconstruir según sea necesario; (iv) asegurar que se brinde a las personas afectadas la oportunidad de mejorar sus ingresos, las actividades generadoras de ingresos y los niveles de vida que tenían antes de que el proyecto las afectara.

Plan de Gestión de Impacto Social (PGIS). Un documento formal y un sistema de gestión asociado que delinea las estrategias a adoptar durante las

diversas fases de un emprendimiento (incluido el cierre) a fin de monitorear, rendir informes, evaluar, examinar y responder de manera proactiva al cambio. De algún modo similar a la Declaración de Impacto Social, la idea de base de un **PGIS** consiste en concentrarse en las estrategias de gestión para abordar los impactos, en lugar de ser solo un listado de impactos potenciales.

Plan de recuperación y mejora de medios de subsistencia. Un plan creado como parte de un proceso de reasentamiento para restablecer y mejorar los medios de subsistencia de las personas luego de haber sido afectadas por un reasentamiento o desplazamiento económico.

Planificación del cierre. El proceso de planificación y gestión del emplazamiento de un proyecto con miras a la situación posterior al cierre; en otras palabras, luego de que se cierra la mina o fábrica. La evaluación de impacto social para la introducción de innoavciones realizada de conformidad con las buenas prácticas permite que se incorpore la planificación posterior al cierre al proceso de planificación y sea prevista en una fase inicial de la construcción. Esto es muy importante, especialmente en un proyecto de minería donde la volatilidad de los precios de los recursos afecta la viabilidad de la mina.

Planificación posterior. Véase *Planificación del cierre.*

Población de interés. Grupo exhaustivo de todas las unidades (como individuos, hogares, empresas, centros) elegibles para recibir una intervención o un tratamiento, y para los cuales una evaluación de impacto se propone estimar los impactos del programa.

Política. Documento elaborado por una organización y que es una declaración de principios, o una declaración general de metas o de fases procedimentales acerca de algún aspecto de importancia organizacional.

Políticas de salvaguardia. Serie de políticas del Banco Mundial que estipulan el desempeño ambiental y social.

Políticas de salvaguardia del Banco Mundial. Serie de 10 políticas de salvaguardia identificadas por el Banco Mundial como importantes a la hora de garantizar que sus operaciones no perjudiquen a las personas o al medio ambiente: Evaluación ambiental (OP 4.01); Zonas en disputa

(OP 7.60); Bosques (OP 4.36); Pueblos indígenas (OP 4.10); Cursos de agua internacionales (OP 7.50); Reasentamiento involuntario (OP 4.12); Hábitats naturales (OP 4.04); Control de plagas (OP 4.09); Recursos culturales físicos (OP 4.11); y Seguridad de las presas (OP 4.37).

Políticas de salvaguardia social del Banco Mundial. Un subgrupo de políticas de salvaguardia (véase arriba) específicas sobre cuestiones sociales (o que las gestiona el Departamento de Desarrollo Social del Banco) y que comprende: Pueblos indígenas (OP 4.10) y Reasentamiento involuntario (OP 4.12).

Potencia (o potencia estadística). Probabilidad de que una evaluación de impacto detecte un impacto (es decir, una diferencia entre el grupo de tratamiento y el grupo de comparación) cuando, de hecho, hay un impacto. La potencia es igual a 1 menos la probabilidad de un error de tipo II, que oscila entre 0 y 1. Los niveles habituales de potencia son 0,8 y 0,9. Los niveles altos de potencia son más conservadores, lo que significa que hay una baja probabilidad de no detectar los impactos reales del programa.

Potencia estadística. La potencia de una prueba estadística es la probabilidad de que la prueba rechace la hipótesis nula cuando la hipótesis alternativa es verdadera (es decir, que no se cometerá un error de tipo II). A medida que la potencia aumenta, la probabilidad de un error de tipo II disminuye. La probabilidad de un error de tipo II se denomina tasa negativa falsa (β). Por lo tanto, la potencia es igual a $1 - \beta$.

Práctica avanzada. En realidad es lo mismo que las mejores prácticas, pero quizás con la idea de que lo que es avanzado siempre cambia, y por consiguiente es un concepto relativo. Tiende a hacer referencia a una actitud de una empresa (que sea una organización que emplee prácticas avanzadas) más que a una propiedad de la práctica.

Predicción. Tarea de identificar posibles impactos futuros para una intervención planeada.

Principios del Ecuador. Marco de principios en materia de sostenibilidad y responsabilidad social corporativa en el sector financiero global. Más específicamente, es un marco de gestión de riesgo que entidades financieras (esto es, bancos) han adoptado para determinar, evaluar y gestionar el

riesgo social y ambiental en proyectos en cualquier parte del mundo y para todos los sectores industriales. Su principal propósito es brindar un estándar mínimo para que se realicen procedimientos de debida diligencia en la toma de decisiones responsable sobre riesgos. Los bancos que han adherido a estos principios se comprometen a implementarlos en sus normas ambientales y sociales internas, procedimientos y estándares para financiar proyectos y acuerdan "*no ofrecer financiación de proyectos ni préstamos corporativos vinculados a proyectos¡ cuando el cliente no cumpla, o no tenga la capacidad de cumplir los Principios de Ecuador*". En esencia, los Principios de Ecuador constituyen una serie de principios de alto nivel; para obtener los lineamientos operativos, los Principios de Ecuador requieren que se cumplan las Normas de Desempeño de la **CFI.**

Principios Rectores. Tiende a hacer referencia a los Principios Rectores de las Naciones Unidas sobre las Empresas y los Derechos Humanos.

Principios voluntarios. En el marco de una evaluación de impacto social para la introducción de innovaciones, generalmente refieren a los Principios voluntarios de seguridad y derechos humanos.

Proceso de cambio social. Un proceso de cambio identificable en comunidades afectadas por proyectos, que es creado, iniciado, habilitado, facilitado y/o exacerbado por medio de una intervención planeada. El proceso de cambio social no es en sí mismo un impacto social, pero podría o no conducir a una situación de impactos sociales, lo cual dependerá del contexto local. Por ejemplo, la inmigración y el reasentamiento son procesos de cambio social que pueden o no provocar impactos sociales.

Proceso de concesión. El proceso mediante el cual un operador obtiene una licencia o permiso en materia de licencias medioambiental, y el proceso por el cual el organismo regulador determina si debe otorgarse.

Proceso de obtención de permiso. Proceso regulatorio que consiste en evaluar y aprobar un proyecto. (Véase también *Licencia ambiental*)

Proceso de resolución de controversias. Un concepto relacionado con acuerdos y contratos (y que con frecuencia constituye una cláusula de ellos), ligado a lo que sucederá si se produce un desacuerdo acerca de la interpre-

tación del contrato/acuerdo o un desacuerdo entre las partes del acuerdo. Es diferente de un mecanismo de reclamos.

Producto. Productos, bienes y servicios tangibles producidos (suministrados) directamente por las actividades de un programa. La generación de productos está directamente bajo el control del organismo ejecutor del programa. El uso de los productos por parte de los beneficiarios contribuye a cambios en los resultados. En evaluación, lógica de programas y gestión de proyectos, el programa entrega productos. El concepto que importa es que los productos rara vez son los resultados previstos, únicamente constituyen un paso hacia alcanzarlos.

Profanización. Proceso por el cual objetos sagrados pasan a ser ordinarios (profanos). La exposición a otras culturas, y especialmente la venta de objetos culturales, los hace perder su valor sagrado.

Programa. Plan coherente y organizado o conjunto ordenado de compromisos, propuestas, instrumentos y/o actividades que describe e implementa una política y que, a la larga, enmarcará a varios proyectos.

Promoción aleatoria. Método de variables instrumentales para estimar los impactos de un programa. El método asigna de forma aleatoria a un subgrupo de unidades una promoción, o incentivo, para participar en el programa. La promoción aleatoria busca aumentar la participación voluntaria en un programa en una submuestra de la mpoblación seleccionada aleatoriamente. La promoción puede adoptar la forma de un incentivo, estímulo o información adicional que motiva a las unidades a inscribirse en el programa, sin influir directamente en el resultado de interés. De esta manera, el programa puede quedar abierto a todas las unidades elegibles.

Propensión por coincidencia de puntajes (PSM. *Propensity Score Matching***)**. Método de coincidencia (*matching*) o pareamiento que depende de los puntajes de propensión para encontrar el mejor grupo de comparación posible para un determinado grupo de tratamiento.

Propensión de puntaje. En el contexto de la evaluación de impacto, la propensión de puntaje es la probabilidad de que una unidad participe en el programa sobre la base de las características observables. Esta puntuación

es un número real entre 0 y 1 que resume la influencia de todas las características observables en la probabilidad de inscribirse en el programa.

Propietarios tradicionales. Un término que se utiliza en Australia para designar los pueblos indígenas que tienen el reclamo válido de ser considerados dueños de la tierra en virtud de un título autóctono.

Proponente. En el marco de una evaluación de impacto social (EIS), las proyecciones refieren a las estimaciones, extrapolaciones, predicciones o pronósticos sobre el estado futuro de una o más variables del impacto social en consideración**.**

Proyección. En el marco de una evaluación de impacto social (EIS), las proyecciones refieren a las estimaciones, extrapolaciones, predicciones o pronósticos sobre el estado futuro de una o más variables del impacto social en consideración.

Proyecto (Project). Un emprendimiento de inversión propuesto que por lo general conlleva la planificación, el diseño y la implementación de actividades específicas.

Proyecto totalmente nuevo. ("*Terreno verde*") Proyecto nuevo, que tiene lugar donde no hubo desarrollo previo; dicho de otro modo, que no implica cuestiones heredadas.

Proyectos lineales. Proyectos como oleoductos o gasoductos, autopistas, ferrocarriles, líneas de transmisión y sistemas de riego que afectan una franja angosta de terreno pero se extienden por muchos kilómetros.

Prueba de placebo. Prueba falsificada que se utiliza para evaluar si los supuestos de un método se mantienen. Por ejemplo, cuando se aplica el método de diferencias en diferencias, se puede implementar una prueba de placebo utilizando un grupo de tratamiento falso o un resultado falso, es decir: un grupo o resultado que se sabe que no se ve afectado por el programa. Las pruebas de placebo no pueden confirmar que los supuestos sean válidos, pero pueden poner de manifiesto los casos en que los supuestos no se sostienen.

Prueba de significancia. Prueba de si la hipótesis alternativa alcanza el nivel predeterminado de significancia con el fin de que esta se acepte de preferencia a la hipótesis nula. Si una prueba de significancia da un valor

p menor que el nivel de significancia estadística (α), la hipótesis nula es rechazada.

Pueblo indígena. A grandes rasgos se lo define como un grupo social y cultural distinto y que manifiesta en alguna medida las siguientes características: autoidentificación como miembro de un grupo cultural distinto y reconocimiento de esta identidad por otros; vínculo colectivo con un hábitat determinado geográficamente o territorio ancestral y a los recursos naturales que allí se encuentran; instituciones culturales, económicas social y/o políticas habituales diferentes de las de la sociedad o cultura dominante; una lengua que a menudo difiere de la lengua oficial del país o de la región.

Punto de referencia. Norma de comparación. Para cada indicador social seleccionado para hacer monitoreo, se debería identificar algún estándar o valor que pueda emplearse como referencia. Por ejemplo, una referencia podrá ser los niveles aceptables para la **OMS** en materia de contaminación del aire o exposición al ruido, o la cantidad esperada de médicos por mil habitantes.

PYMES (SME). Pequeñas a medianas empresas, su definición varía de un país a otro según los criterios sobre la cantidad de trabajadores o los ingresos anuales. Cuando se trata exclusivamente de empresas muy pequeñas o unipersonales, *microempresa* es el término que se emplea. Las PYMES son importantes en una evaluación de impacto social para la introducción de innovaciones, ya que a menudo están dispersas, desorganizadas y con frecuencia se ven afectadas por los proyectos. Con la debida consideración, también pueden tener una gran participación activa a través de un compromiso a realizar las adquisiciones en el ámbito local.

Reasentamiento. El proceso planificado de reubicar personas y comunidades de un lugar a otro como parte de la adquisición de tierras, inducida por el proyecto, necesaria para que se pueda ejecutar un proyecto. El reasentamiento se considera involuntario cuando el lugar del proyecto es fijo y las comunidades locales realmente no tienen otra opción que ser reubicados; mientras que el reasentamiento se considera voluntario cuando el Estado no ejerce ni amenaza con ejercer el derecho de expropiación, ni se percibe

una tal amenaza, y los individuos afectados tienen una oportunidad real con respecto a si aceptan o no su reasentamiento.

Reasentamiento involuntario. El reasentamiento de personas en razón de un proyecto, en situaciones en las que el Estado ha ejercido o amenazado con ejercer su facultad de dominio eminente. Véase también *Reasentamiento*)

Reasentamiento voluntario. Véase *Reasentamiento.*

Recién llegados. Residentes que llegan a instalarse en un determinado lugar.

Regalía. En términos generales, una regalía es lo que paga una parte (el licenciatario) a otra parte (el licenciante) por el derecho de uso continuado de un bien. En el marco de una evaluación de impacto social para la introducción de innovacionesl as regalías son pagos que las empresas dedicadas a la extracción de recursos realizan a los gobiernos y/o a los propietarios Tradicionales de las tierras para acceder a los recursos que se extraen.

Registro de compromisos. Un documento público formal que registra toda declaración o promesa de la empresa a la comunidad especialmente en relación con cualquier beneficio o mitigación prometido.

Rehabilitación. En la evaluación de impactos, hace referencia en general a retrotraer el paisaje al estado en que se encontraba antes del proyecto (*restauración*), o cuando esto no es posible, al menos hacer que el paisaje resulte aceptable para la gente (*recuperación*). En minería, por ejemplo, se refiere al reemplazo de las capas del suelo y a la revegetación de la tierra.

Residentes de larga data. Hace referencia a los que residen desde hace tiempo en un lugar, se emplea en contraste a los recién llegados. Los intereses de los veteranos a menudo difieren de los de los nuevos.

Resiliencia. Capacidad de una comunidad para recuperarse de los impactos que la amenazan.

Resistencia. Capacidad de una comunidad para resistir el cambio, ser capaz de resistirse frente a un proyecto de desarrollo injustificado.

Resolución/gestión de conflictos. Las intervenciones a menudo modifican las relaciones de poder entre los diferentes grupos de la sociedad. Algunos grupos pueden perder mientras otros ganan, y como consecuencia pueden

surgir conflictos. Los conflictos son un componente normal de la interacción social, pero cuando se vuelven problemáticos ejercen un impacto negativo sobre todas las partes intervinientes. Por consiguiente, se hace necesario contar con mecanismos y técnicas eficaces para la prevención, manejo y resolución de los conflictos, de modo de solucionarlos o de mantenerlos dentro de límites aceptables. La transparencia y el intercambio de información pueden contribuir a eliminar conflictos provocados por un conocimiento incompleto o distorsionado. La aceptación y la implementación de un amplio espacio para la expresión de diferentes puntos de vista pueden ayudar a prevenir el desarrollo de formas más destructivas de conflicto.

Restauración del paisaje. Véase *Rehabilitación.*

Resultado. Resultado de interés que se mide a nivel de los beneficiarios del programa. Resultados que deben alcanzarse una vez que la población beneficiaria utilice los productos del proyecto. Los resultados no están directamente bajo el control de un organismo ejecutor del programa. En ellos influye tanto la implementación de un programa (las actividades y productos que genera) como las respuestas de las conductas de los beneficiarios expuestos a ese programa (el uso que los beneficiarios hacen de los beneficios a los que están expuestos). Un resultado puede ser intermedio o final (de largo plazo). Los resultados finales son resultados más distantes. La distancia se puede interpretar en términos de tiempo (se tarda más en conseguir el resultado) o en términos de causalidad (se requieren numerosos vínculos causales para alcanzar el resultado y en ello influyen múltiples factores). Lo que se espera que logre una intervención específica en el ámbito del desarrollo.

Retorno social sobre la inversión. Metodología para calcular el beneficio adicional más allá del retorno de inversión financiero inmediato en proyectos o actividades, por ejemplo, de la financiación en inversión social. Generalmente se expresa como proporción relativa a los recursos invertidos.

Riesgo. Puede significar la probabilidad de que se produzca algún acontecimiento, pero el término puede emplearse con un significado levemente diferente para hacer referencia a un suceso incierto (o probabilidad desco-

nocida) que, si ocurre, afectará la concreción de uno o más objetivos. A menudo se crean subcategorías de riesgos. Por ejemplo, los riesgos no técnicos (o riesgos sociales) se relacionan con cuestiones administrativas, jurídicas, sociales y políticas de un proyecto; mientras que los riesgos técnicos son los aspectos físicos, estructurales, ingenieriles y medioambientales.

Riesgo reputacional. El riesgo potencial para la reputación de una organización, al vincularse o ser vinculada con una práctica en particular.

Riesgo social. El Banco Mundial define el riesgo social como "*la posibilidad de que la intervención cree, refuerce o profundice la desigualdad o el conflicto social o ambos, o de que las actitudes y acciones de los principales interesados puedan subvertir el logro del objetivo de desarrollo o de que los principales interesados no asuman la responsabilidad por el logro del objetivo de desarrollo ni cuenten con los medios para lograrlo. Tales riesgos pueden surgir del contexto sociocultural, político, operativo o institucional*".

Riesgos no técnicos. Están relacionados con la gestión y los aspectos jurídicos, sociales y políticos de un proyecto. En el sector industrial, a veces se los denomina riesgos *externos* ya que ocurren como resultado de circunstancias ajenas al control de los directores de proyectos.

Rito/ritual. Acto ceremonial que es expresión de cultura y comunidad.

Saber y hacer saber. Una frase popularizada por John Ruggie. En lugar de ser vulnerables y quedar expuestas a las campañas de denuncia y descrédito, las empresas deberían demostrar que han internalizado su respeto por los derechos humanos, a través de su compromiso con los procesos de debida diligencia.

Sabotaje. Un acto deliberado destinado a perjudicar a una empresa por medio de obstrucción, perturbación o destrucción. Las partes interesadas damnificadas podrían ejercer una acción directa contra un proyecto bloqueándolo o destruyendo equipos ya sea en un intento por atraer publicidad, retrasar el proyecto o simplemente por venganza.

Sagrado. Que tiene significancia espiritual o religiosa.

Salud mental. En pocas palabras, es el nivel de bienestar psicológico y la ausencia de cualquier trastorno mental. Comprende la habilidad de un individuo de disfrutar de la vida, así como su bienestar subjetivo, autoeficacia percibida, autonomía, competencia y desarrollo personal de su potencial intelectual y emocional.

Seguidores. Este término en general refiere a los civiles que siguen ejércitos, que usualmente satisfacen las necesidades de los soldados con bienes y servicios no provistos por las autoridades militares, esto es, determinados alimentos, lavado de ropa, alcohol y drogas, cuidado de enfermos y servicios sexuales. En su empleo actual más generalizado describe a aquellos empresarios que acosan de servicios a los obreros en campamentos de construcción y terrenos de proyectos. (Véase *Efecto tarro de miel*)

Seguridad socialmente responsable. Forma de brindar servicios de seguridad en torno a un proyecto atendiendo a los derechos humanos y otras cuestiones sociales, y que probablemente esté en consonancia con los Principios voluntarios de seguridad y derechos humanos.

Selección. Se produce cuando la participación en el programa se basa en las preferen cias, decisiones o características no observables de los participantes o de los administradores del programa.

Sensibilidad cultural. Cualidad individual de tener conciencia de diferencias culturales y de saber cómo manejarse en situaciones multiculturales. Muchos impactos sociales surgen por falta de sensibilidad cultural de gran parte del personal del proyecto.

Sentido comunitario. El grado de noción que los individuos de un determinado lugar tienen de formar parte de una comunidad, y de ayudar a través de la participación en actividades comunitarias y de ser buen vecino (esto es, ejercer la buena vecindad).

Sentido de comunidad. Véase *Sentido comunitario.*

Sentido de pertenencia. Una relación personal de un individuo con su entorno local, tanto social como natural, que mantiene en su vida cotidiana.

Servicios del ecosistema. La noción de que el medio ambiente (un ecosistema) provee una variedad de servicios (y productos) de los cuales dependen los seres humanos. Por lo general, estos servicios se identifican como: apro-

visionamiento (por ejemplo, la producción de alimentos y agua); regulación (por ejemplo, el control del clima y de las enfermedades); secundarios (por ejemplo, los ciclos de los nutrientes y la polinización de los cultivos); y culturales (por ejemplo, beneficios espirituales y recreativos). Para ayudar a informar a los responsables de la toma de decisiones, frecuentemente se asigna a los servicios del ecosistema un valor económico.

Sesgo. En la evaluación de impacto, el sesgo es la diferencia entre el impacto que se calcula y el verdadero impacto del programa.

Sesgo de cobertura. Se produce cuando un marco muestral no coincide exactamente con la población de interés.

Sesgo de selección. El impacto estimado sufre un sesgo de selección cuando se desvía del impacto verdadero en presencia de la selección. Esto suele ocurrir cuando se correlacionan motivos no observados para participar en el programa con los resultados. Este sesgo normalmente acontece cuando el grupo de comparación es no elegible o se autoexcluye del tratamiento.

Sesgo de sustitución. Efecto no intencionado de la conducta que afecta al grupo de comparación. Las unidades que no fueron seleccionadas para recibir el programa pueden encontrar buenos sustitutos para el tratamiento a través de su propia iniciativa.

Significancia. La significancia estadística señala la probabilidad de cometer un error de tipo I; es decir, la probabilidad de detectar un impacto que en realidad no existe. El nivel de significancia suele señalarse con el símbolo griego α (alfa). Los niveles más habituales de significancia son del 10%, 5% y 1%. Cuanto menor sea el nivel de significancia, mayor será la confianza de que el impacto estimado es real. Por ejemplo, si el nivel de significancia se fija en 5%, se puede tener un 95% de confianza al concluir que el programa ha tenido impacto, si de hecho se observa un impacto significativo.

Simulaciones ex ante. Evaluaciones que utilizan datos disponibles para simular los efectos previstos de un programa o de la reforma de una política en los resultados de interés.

Sistema de gestión ambiental. Serie de actividades continuas y planeadas, basadas en el concepto de mejora continua, que una empresa lleva a cabo

para gestionar mejor sus impactos ambientales. Está estipulada por la norma ISO 14001.
Sistema de gestión social. Sistema de gestión que específicamente aborda los problemas sociales en una empresa o proyecto.
Sistema de tenencia de tierras. Las disposiciones legales por las cuales se establecen formalmente la propiedad de la tierra y todo proceso para la transferencia intergeneracional o venta de la propiedad de la tierra.
Sitio sagrado. Un sitio (lugar en un paisaje) con un significado espiritual especial para la población local. Aunque se lo asocia habitualmente con los pueblos indígenas, podría aplicarse en un sentido más amplio para referirse a otros lugares y santuarios espirituales y religiosos con una gran significación de patrimonio cultural.
Situación de calle. La situación personal de no tener una vivienda, de no tener un lugar permanente donde estar. En un nivel, puede estar referido a personas sin domicilio fijo que dependen de los servicios de vivienda de emergencia y que de lo contrario dormirían en las calles (es decir, *dormir a la intemperie*); en otro nivel, puede significar un sentimiento de alienación tal que aunque pueda haber un lugar de residencia físicamente adecuado, otros elementos emocionales ligados a él significan que no se lo percibe como un hogar.
Sociedad civil. La red de individuos y grupos (tanto formal como informal) –y sus conexiones, prácticas y normas sociales– que comprenden las actividades de una sociedad independientemente de las instituciones estatales y del mercado. Incluye organizaciones religiosas, grupos comunitarios, fundaciones, gremios, asociaciones profesionales, sindicatos, instituciones académicas, medios de comunicación, grupos de presión o activistas, partidos políticos, etc.
Subsistencia. Este concepto se encuentra en términos como economía de subsistencia, medios de vida para subsistencia, agricultura de subsistencia, pesca de subsistencia y minería de subsistencia (esto es, minería artesanal). Se refiere a la economía informal o no mercantil (en vez de la economía monetaria) en la cual las personas producen sus propios bienes y servicios o los intercambian en trueques y no por efectivo.

Supuesto de estabilidad del valor de la unidad de tratamiento (SUTVA). Requisito básico de que el resultado de una unidad no debería verse afectado por la asignación del tratamiento a otras unidades. Esto es necesario para asegurar que la asignación aleatoria produzca estimaciones de impacto no sesgadas.

Tabú. Algo que está culturalmente prohibido. Los proyectos a menudo transgreden tabúes por desconocimiento, generando mucho agravio.

Tamaño del efecto. Magnitud del cambio en un resultado, que es causado por una intervención.

Tecnocrático. Un término peyorativo que hace referencia a personas e instituciones que poseen excelentes conocimientos técnicos pero carecen de conciencia social y comprensión social; y en particular que toman decisiones sobre la base de su conocimiento técnico, sin prestar la debida atención al contexto social y político.

Teoría del cambio. Explica los canales a través de los cuales los programas pueden influir en los resultados finales. Describe la lógica causal de cómo y por qué un programa, una modalidad de programa o una innovación de diseño en particular logrará sus resultados deseados. Una teoría del cambio es una pieza clave en cualquier evaluación de impacto, dada la focalización de causa y efecto de la investigación.

Titular de derechos. Individuos y grupos cuyos derechos han sufrido un impacto. Podría incluir a todas las partes interesadas. El término es en realidad similar a *parte interesada* pero el empleo de *titular de derechos* implica una conexión con el enfoque basado en los derechos humanos y una conciencia precisa de que estas personas podrían tener capacidad jurídica.

Tratamiento. Véase *Intervención*.

Tratamiento en los tratados (**TOT**, por sus siglas en inglés). Las estimaciones **TOT** miden la diferencia en los resultados entre las unidades que en efecto reciben el tratamiento y el grupo de comparación.

Unidad. Persona, hogar, comunidad, empresa, escuela, hospital u otra unidad de observación que pueda ser objeto de un programa o verse afectada por él.

Validez externa. Una evaluación es externamente válida si la muestra de la evaluación representa con precisión a la población de unidades elegibles. Los resultados de la evaluación, posteriormente, se pueden generalizar a la población de unidades elegibles. Estadísticamente, para que una evaluación de impacto sea externamente válida, la muestra de la evaluación debe ser representativa de la población de interés. Véase también *Validez interna*.
Validez interna. Una evaluación es internamente válida si proporciona una estimación precisa del contrafactual mediante un grupo de comparación válido.
Valor compartido. Una manera de pensar acerca del rol de una empresa, que reconoce que son las necesidades de la sociedad, y no solo las necesidades económicas convencionales, las que definen los mercados, y que el objetivo de la empresa debe definirse como la creación de valor compartido y no solo utilidades, de modo que la sociedad se beneficie tanto como la empresa. Esta visión también contempla que los perjuicios sociales generan frecuentemente costos para las empresas, en forma de riesgos sociales, y por consiguiente deben manejarse cuidadosamente**.**
Valor intrínseco. Concepto filosófico según el cual un objeto o entidad (como la naturaleza, un lugar determinado, una especie rara) tiene un valor o calidad inherente más allá de su valor de uso o instrumental para los humanos.
Valores. Suposiciones abstractas y a menudo subconscientes que los individuos hacen acerca de lo que está bien o tiene importancia en sus vidas. Por lo general, se organizan en un sistema de valores. Los valores y los sistemas de valores pueden variar sustancialmente de un grupo cultural a otro.
Variable. En la terminología estadística, se trata de un símbolo que representa un valor que puede variar.
Variable dependiente. Normalmente, es la variable de resultado. Se trata de la variable que hay que explicar, por oposición a las variables explicativas.

Variable explicativa. También conocida como variable “independiente”. Se trata de una variable utilizada en el lado derecho de una regresión para ayudar a explicar la variable dependiente en el lado izquierdo de la regresión.
Variable instrumental (VI). También conocida como instrumento. Se basa en el uso de una fuente externa de variación para determinar la probabilidad de participación en el programa cuando la participación en el mismo está relacionada con los resultados potenciales. El instrumento se encuentra fuera del control de los participantes y no tiene relación con las características de los mismos.
Variables no observadas. Se trata de características no observables. Pueden incluir particularidades como la motivación, las preferencias u otros rasgos de la personalidad que son difíciles de medir.
Vías causales. Un concepto vinculado con el de pensamiento sistémico y el de evaluación. Refiere a las relaciones causales (al menos a las correlaciones) entre diversos elementos dentro de un sistema. En una evaluación de impacto social para introducción de innovacioens, refieren a las secuencias de las experiencias de impactos primarios y secundarios y entre los procesos de cambio social y los impactos sociales.
Vías o cadenas de impacto. Se refiere a los vínculos entre los impactos primarios (impactos de primer orden) y los impactos secundarios; así como a los vínculos entre los procesos de cambio social tal como la inmigración.
Visión comunitaria. Proceso de crear consenso sobre el futuro que la comunidad quiere, y luego decidir sobre lo que sea necesario para alcanzarlo. Es tanto el proceso de creación de una visión como el producto de esa visión.
Viviendas de fin de semana. Moradas de personas quienes durante los fines de semana y vacaciones se trasladan a sus segundos hogares. A veces el término también se usa para referirse a las casas. A menudo son un elemento importante en las evaluaciones de impacto social de introducción de innovaciones, pero acceder a ellos es difícil. Algunas veces los proyectos los afectan adversamente pero sus problemas son diferentes de los del resto de la comunidad afectada.

Vulnerabilidad. Una situación o estado caracterizado por una baja resistencia o un mayor riesgo y la capacidad reducida de un individuo, grupo o comunidad de hacer frente a conmociones o impactos adversos. La vulnerabilidad está asociada con una situación socioeconómica baja, incapacidad, etnicidad o uno o más de los muchos factores que influyen en la capacidad de las personas para acceder a los recursos y a las oportunidades de desarrollo.

Zona de influencia o del proyecto. Comprende: el(los) emplazamiento(s) principal(es) del proyecto e instalaciones relacionadas que un proponente (o sus contratistas) desarrolla o controla; instalaciones asociadas construidas como resultado del proyecto (incluso si no fueran financiadas directamente por el proyecto sino por un cliente o un tercero, que puede ser el gobierno), y cuya viabilidad y existencia depende exclusivamente del proyecto y cuyos bienes y servicios son esenciales para la operación exitosa del proyecto; áreas potencialmente afectadas por impactos acumulativos de otro emprendimiento planeado del proyecto; y áreas potencialmente afectadas por impactos que surgen de acontecimientos no planeados pero predecibles causados por el proyecto y que pueden ocurrir posteriormente o en otro emplazamiento.

Zona de influencia primaria. Se refiere a los impactos sociales que ocurren en una zona de influencia primaria por la medida propuesta y que tienen lugar en el mismo momento y lugar.

Bibliografía

AIEI (2015). *Evaluación de Impacto Social: Lineamientos para la evaluación y gestión de impactos sociales de poryectos*. Asociación Internacional para la Evaluación de Impactos. Banco Interamericano de desarrollo, en: https://www.iaia.org/pdf/Evaluacion-Impacto-Social-Lineamientos.pdf

Angelucci, M. y Di Maro, V. (2015). Programme Evaluation and Spillover Effects. *Journal of Development Effectiveness* (doi: 10.1080 /19439342.2015.1033441).

Banerjee, A.; E. Duflo. (2009). The Experimental Approach to Development Economics.*Annual Review of Economics 1: 151–78.*

Bernal, R. y Peña X. (2011). *Guía práctica para la evaluación de impacto.* Perú: Universidad de los Andes: Facultad de Economía.

CGIB (2019). *Pesquisa sobre o uso das Tecnologias de Informação e Comunicação nas escolas brasileiras – TIC. Educação 2018.*Comité Gestor de Internet en Brasil, São Paulo, en: https://www.cgi.br/publicacao/pesquisa-sobre-o-uso-das-tecnologias-de-informacao-e-comunicacao-nas-escolas-brasileiras-tic-educacao-2018/

CONEVAL (2020). *La política social en el context de la pandemia por el virus SARS-CoV-2 (COVID-19) en México.* México: Consejo Nacional de Evaluación de la Política de Desarrollo Social, en: https://www.coneval.org.mx/Evaluacion/IEPSM/Documents/Politica_Social_COVID-19.pdf

DRALE (2020). *Diccionario de la Real Academia Española*, en: https://www.rae.es/

DOF (2020). *Decreto por el que se establecen las medidas de austeridad que deberán observar las dependencias y entidades de la Administra-*

ción Pública Federal bajo los criterios que en el mismo se indican, Diario Oficial de la Federación, en: https://www.dof.gob.mx/nota_detalle.php?codigo=5592205&fecha=23/04/2020

Duflo, E.; Glennerster, R.; y Kremer, M.(2008). *Using Randomization in Development Economics Research: A Toolkit.* In *Handbook of Development Economics,* vol. 4, ed. T. Paul Schultz and John Strauss, 3895–962. Amsterdam: North-Holland.

FAO/PMA (2019). *Fortaleciendo los programas de alimentación escolar: el trabajo conjunto de FAO y WFP en América Latina y el Caribe.* Organización de las Naciones Unidas para la Alimentación y la Agricultura/Programa Mundial de Alimentos. Panamá, en: https://es.wfp.org/publicaciones/fortaleciendo-los-programas-de-alimentacion-escolar-el-trabajo-conjunto-de-fao-y-wfp

Gertler, P. J.; Martínez, S.; Premand, P.; Rwlings, L. B. y Vermeersch, Ch. M. J. (2017). *La evaluación de impacto en la práctica 2a. ed*. Washington, D.C.: Banco Interamericano de Desarrollo y Banco Mundial.

Heckman, J. J.; Ichimura, H. y Todd, P. (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *Review of Economic Studies 64* (4): 605–54.

Heckman, J. J. y Robb, R. (1985). *Alternative Methods for Estimating the Impact of Interventions*.In Longitudinal Analysis of Labor Market Data, ed. James Heckman and Burton Singer, 156–245. New York: Cambridge University Press.

Heckman, J. J. y Vytlacil, E. J. (2000). *Local Instrumental Variables. NBER Technical Working Paper 252*. National Bureau of Economic Research, Cambridge, M A.

Heckman, J. J. y Vytlacil, E. (2005). Structural Equations, Treatment Effects, and Econometric Policy Evaluation. *Econometrica 73 (3):* 669–738.

Imbens, G. y Angrist, J. (1994). Identifi cation and Estimation of Local Average Treatment Effects. *Econometrica 62 (2)*: 467–76.

Khandker, S. R. (2006). Microfinance and Poverty: Evidence Using Panel Data from Bangladesh.*World Bank Economic Review 19* (2): 263–86.

Khandker, S. R.; Koolwal, G. B. y Samad, H. A. (2017). *The Handbook on Impact Evaluation. Quantitative Methods and Practices.* Washington, D. C.: The Worldbank.

King, E. M. y Behrman. J. R. (2009). Timing and Duration of Exposure in Evaluations of Social Programs. *World Bank Research Observer* 24 (1): 55–82.

Lee, D. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *Review of Economic Studies* 76 (3):1071–1102.

Mejía-Trejo, J. (2011). *Estudios del Futuro Tecnológico: definiciones hacia un modelo conceptual de prospectiva*. México: CUCEA, en: http://www.cucea.udg.mx/sites/default/files/documentos/adjuntos_pagina/estudios_del_futuro_tecnologico_definiciones_hacia_un_modelo_conceptual_de_prospectiva_0.pdf

Mejía-Trejo, J. (2019a). *Técnicas Dependientes con SPSS*, en: https://buk.com.mx/9786079878207/description

Mejía-Trejo, J. (2019b). *Diseño de Cuestionarios y Creación de Escalas. Uso del EQS en las ciencias económico administrativas,* en: https://buk.com.mx/9786075384672/description

Moffitt, R. (2008). Estimating Marginal Treatment Effects in Heterogeneous Populations. Economic Working Paper Archive 539, Johns Hopkins University, Baltimore, MD, en: http://www.econ.jhu.edu/people/moffi tt/welfl s0_v4b.pdf

OIT (2020). *Covid-19 y el mundo del trabajo: repercusiones y respuestas*. Organización Internacional del Trabajo, en: https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/documents/briefingnote/wcms_739158.pdf.

OMS (2017). *Global Health Expenditure*. Organización Mundial de la Salud, en: https://apps.who.int/nha/database/Select/Indicators/en.

ONU CEPAL (2015). *La Agenda 2030 y los Objetivos de Desarrollo Sostenible Una oportunidad para América Latina y el Caribe.* Santiago,

Chile: ONU, Comisión Económica para América Latina y el Caribe, en: https://www.un.org/sustainabledevelopment/es/

ONU CEPAL (2019). *Panorama Social de América Latina*. ONU, Comisión Económica para América Latina y el Caribe, en: https://www.cepal.org/es/publicaciones/44969-panorama-social-america-latina-2019

ONU CEPAL (2020a). *Informe Especial COVID-19 América Latina y el Caribeante la pandemia del COVID-19. Efectos Económicos y Sociales*. ONU, Comisión Económica para América Latina y el Caribe, en: https://www.cepal.org/es/publicaciones/45337-america-latina-caribe-la-pandemia-covid-19-efectos-economicos-sociales

ONU CEPAL (2020b). *Dimensionar los efectos del COVID-19 para pensar en la reactivación*. ONU, Comisión Económica para América Latina y el Caribe, en: https://www.cepal.org/es/publicaciones/45445-dimensionar-efectos-covid-19-pensar-la-reactivacion

OPS (2019) *Indicadores básicos 2019: Tendencias de la salud en las Américas*. Organización Panamericana de la Salud, en: https://iris.paho.org/handle/10665.2/51543

OPS (2020) *Actualización epidemiológica: dengue*. Organización Panamericana de la Salud, en: https://www.paho.org/hq/index.php?option=com_docman&view=download&category_slug=dengue-2158&alias=51692-7-de-febrero-de-2020-dengue-actualizacion-epidemiologica-1&Itemid=270&lang=es

PDR (2020). *Nada nos hará regresar al pasado, afirma presidente al presentar informe al pueblo de México*. Comunicado de Prensa 6 de abril de 2020, en: Recuperado de https://www.gob.mx/presidencia/prensa/nada-nos-hara-regresar-al-pasado-afirma-presidente-al-presentar-informe-al-pueblo-de-mexico?idiom=es-MX

Ravallion, M. (2008). Evaluating Anti-Poverty Programs. *In Handbook of Development Economics,* vol. 4, ed. T. Paul Schultz and John Strauss, 3787–846. Amsterdam: North-Holland.

RAO, V. y Woolcock, M. (2003). *Integrating Qualitative and Quantitative-Approaches in Program Evaluation* en: F. J. Bourguignon y L. Pereira

da Silva, *The Impact of Economic Policies on Poverty and Income Distribution: Evaluation Techniques and Tools*: 165–90. Nueva York: Oxford University Press.

Semykina, A. y Wooldridge. J. M. (2005). *Estimating Panel Data Models in the Presence of Endogeneity and Selection: Theory and Application*. Working Paper, Michigan State University, East Lansing, MI.

Todd, P. (2007). *Evaluating Social Programs with Endogenous Program Placement and Selection of the Treated*. In *Handbook of Development Economics*, vol. 4, ed. T. Paul Schultz and John Strauss, 3847–94. Amsterdam: North-Holland.



Evaluación de impacto social
en proyectos de innovación vía STATA
Métodos: Estimación por variable instrumental,
regresión doscontinua, elección de método a usar
Tomo II

Registrado y producido como libro electrónico en mayo de 2021
en los talleres gráficos de TRAUCO Editorial
Camino Real a Colima 285 int. 56
Teléfono: (33) 32.71.33.33
Tlaquepaque, Jalisco.

En el mundo empresarial e industrial, el diseño e implementación de proyectos normalmente toma en cuenta el impacto económico, financiero y hasta el político o ambiental en la introducción de innovaciones. Sin embargo, en los primeros veinte años del siglo XXI, han sucedido diversos acontecimientos que han demostrado que la evaluación de impacto social en proyectos de innovación es de vital importancia. De hecho, en los tiempos de la nueva normalidad que se avizora como la era PosCOVID-19, esto toma particular relevancia dado que todas las políticas y acciones que emitan empresas y gobiernos, deberán contar con el aval necesario de una evaluación de impacto social a la introducción de innovaciones.

Es por esta razón, que la presente obra está orientada a describir tanto a propios como ajenos al tema, lo qué es el impacto social, sus características, condiciones e implicaciones, los principales métodos utilizados para calcularla, que demanda que los recursos y acciones de innovación a diseñar e implementar, reflejen altos estándares de impacto social que fomenten el bienestar, particularmente en los países emergentes.

Para lograrlo, esta obra está dividida en una colección de dos tomos, correspondiendo al Tomo II:

La descripción y aplicación de los métodos de estimación por variable instrumental (Instrumental Variables); regresión discontinua (Regression Discontinuity) así como la elección de la muestra y diseño de recolección de datos; la descripción de lineamientos de gestión de proyectos de evaluación de impacto y su implicación en la época PosCovid-19 describiéndose una variedad de ejemplos de proyectos de innovación. Se anexa al final, un breve manual de la configuración y manejo del software STATA





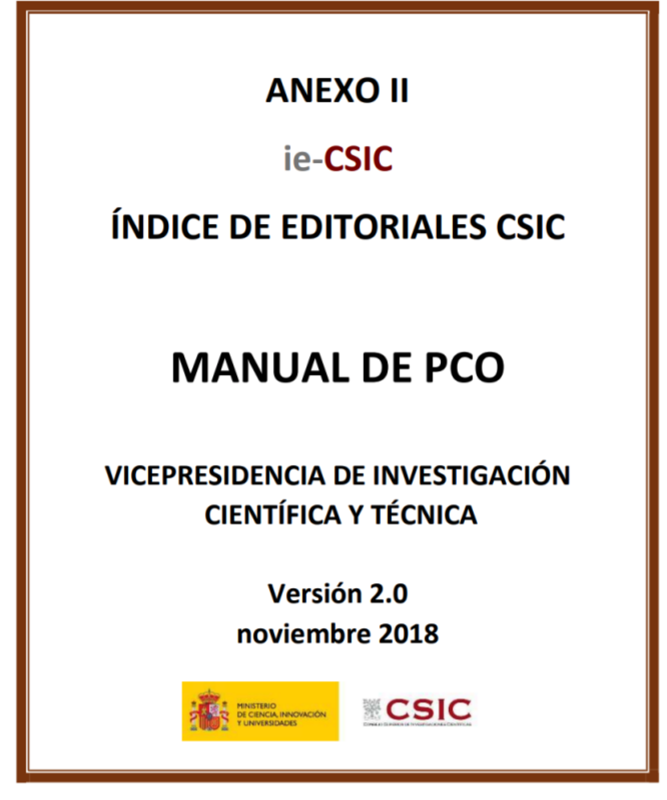


Libro científico que presenta:

1. **Acceso universal al conocimiento a través del:**
   a. **Portal de productividad docente Doctorado en Ciencias de la Administración (DCA) de la Universidad de Guadalajara (UdeG):**
   http://dca.cucea.udg.mx/sites/default/files/adjuntos/2021_tomo_ii.evaluacion_de_impacto_social_en_proyectos_de_innovacion_via_stata.pdf

   b. **Editorial BUK:**
   https://buk.com.mx/9786079908621/description

2. **Reconocimiento del Consejo Superior de investigaciones Científicas (CSIC, España) a la editorial de la Universidad de Guadalajara como nivel Alto.**
   https://glosariobibliotecas.files.wordpress.com/2020/05/ie-csic_2018.pdf



ANEXO II

ie-CSIC

ÍNDICE DE EDITORIALES CSIC

MANUAL DE PCO

VICEPRESIDENCIA DE INVESTIGACIÓN CIENTÍFICA Y TÉCNICA

Versión 2.0
noviembre 2018

Universidad de Extremadura	MEDIO
Universidad de Guadalajara México	ALTO
Universidad de Guanajuato	BAJO
Universidad de Jaén	MEDIO

3. Constancias de dictámenes de la evaluación por pares académicos a doble ciego, a las que fue sometida la obra basada en el Reglamento para Producción Editorial de la Evaluación de la obras, de la Universidad de Guadalajara.

4. Reconocimientos de participación como evaluadores y dictaminadores de la obra.



A quien corresponda:

Por medio de la presente se hace constar que de acuerdo con los registros de esta Secretaría Académica, la obra titulada:

"Evaluación de Impacto Social en proyectos de Innovación vía STATA . Métodos: Estimación por Variable Instrumental, Regresión Discontinua, Elección de método a usar."

ISBN 978-607-571-167-6 TOMO II. COLECCIÓN 978-607-571-165-2

Autor: Juan Mejia Trejo

Se publicó en el año 2021, por contar con los elementos teóricos, metodológicos, técnicos y de redacción de acuerdo con los resultados de la evaluación por pares académicos a doble ciego a la que fue sometida la obra, esto de conformidad en con lo establecido en el Reglamento para la Producción Editorial de este Centro Universitario, en su Título Tercero, De la evaluación de las Obras, Articulo 9, incisos a y b.

Se extiende la presente para los fines que al interesado convenga.

Atentamente
"Piensa y Trabaja"
"Año del Legado de Fray Antonio Alcalde en Guadalajara"
Zapopan, Jal., 10 de diciembre de 2021

Dr. José María Nava Preciado
Secretario Académico

SECRETARÍA ACADÉMICA CENTRO UNIVERSITARIO DE CIENCIAS ECONÓMICO ADMINISTRATIVAS

Periférico Norte 799, Edificio de Rectoría, Planta Alta, Núcleo Universitario Los Belenes C.P. 45100
Zapopan, Jalisco, México Tel. (0133) 3770-3315 y 3770-3300 Ext. 25022, 25024
www.cucea.udg.mx

UNIVERSIDAD DE GUADALAJARA
CENTRO UNIVERSITARIO DE CIENCIAS ECONÓMICO ADMINISTRATIVAS
SECRETARÍA ACADÉMICA

A quien corresponda:

Por medio de la presente, el Centro Universitario de Ciencias Económico Administrativas a través de su Secretaría Académica agradece a:

CARLOS OMAR AGUILAR NAVARRO

Por su participación como evaluador y dictaminador del trabajo de investigación propuesto para su publicación titulado "**Evaluación de Impacto Social en proyectos de Innovación vía STATA . Métodos: Estimación por Variable Instrumental, Regresión Discontinua,Elección de método a usar . TOMO II**", en conformidad con lo establecido en el Reglamento para la Producción Editorial de este Centro Universitario, en su Título Tercero, Articulo 9.

Se extiende la presente para los fines que al interesado convenga.

Atentamente

"Piensa y Trabaja"

"Año de la transición energética de la Universidad de Guadalajara"

Zapopan, Jal., 10 de diciembre de 2020

SECRETARÍA ACADÉMICA
CENTRO UNIVERSITARIO
DE CIENCIAS
ECONÓMICO ADMINISTRATIVAS

Dr. José María Nava Preciado

Secretario Académico

Periférico Norte 799, Edificio de Rectoría, Planta Alta, Núcleo Universitario Los Belenes C.P. 45100
Zapopan, Jalisco, México Tel. (0133) 3770-3315 y 3770-3300 Ext. 25022, 25024

www.cucea.udg.mx

UNIVERSIDAD DE GUADALAJARA
CENTRO UNIVERSITARIO DE CIENCIAS ECONÓMICO ADMINISTRATIVAS
SECRETARÍA ACADÉMICA

A quien corresponda:

Por medio de la presente, el Centro Universitario de Ciencias Económico Administrativas a través de su Secretaría Académica agradece a:

ARIEL VAZQUEZ ELORZA

Por su participación como evaluador y dictaminador del trabajo de investigación propuesto para su publicación titulado "**Evaluación de Impacto Social en proyectos de Innovación vía STATA . Métodos: Estimación por Variable Instrumental, Regresión Discontinua,Elección de método a usar . TOMO II**", en conformidad con lo establecido en el Reglamento para la Producción Editorial de este Centro Universitario, en su Título Tercero, Articulo 9.

Se extiende la presente para los fines que al interesado convenga.

Atentamente

"Piensa y Trabaja"

"Año de la transición energética de la Universidad de Guadalajara"

Zapopan, Jal., 10 de diciembre de 2020

SECRETARÍA ACADÉMICA
CENTRO UNIVERSITARIO
DE CIENCIAS
ECONÓMICO ADMINISTRATIVAS

Dr. José María Nava Preciado

Secretario Académico

Periférico Norte 799, Edificio de Rectoría, Planta Alta, Núcleo Universitario Los Belenes C.P. 45100
Zapopan, Jalisco, México Tel. (0133) 3770-3315 y 3770-3300 Ext. 25022, 25024

www.cucea.udg.mx